Multi-Class Bayesian Segmentation of Ultrasound Images to Enable Safe Needle Insertion in Femoral Vessels

Edward Chen[1] EDWARDC2@ANDREW.CMU.EDU
FNU Abhimayu[1] ABHIMAN2@ANDREW.CMU.EDU
Ananya Bal[1] ABAL@ANDREW.CMU.EDU
Nishanth Thumbavanam Arun[1] NTHUMBAV@ANDREW.CMU.EDU
Andrew Schoenling[2] SCHOENLINGAJ@UPMC.EDU
Joo Yoon[2] YOONJH@UPMC.EDU
Lenny Weiss[3] WEISSLS2@UPMC.EDU
Kelvin Kwofie[3] KWOFIEKZ@UPMC.EDU
Marek Radomski[3] RADOMSKIMA@UPMC.EDU
Frank Guyette[3] GUYEFX@UPMC.EDU
Michael Pinsky [2] PINSKY@PITT.EDU
Howie Choset[1] CHOSET@CMU.EDU
John Galeotti[1] JGALEOTTI@CMU.EDU

[1] *School of Computer Science, Carnegie Mellon University, Pittsburgh*
[2] *Department of Critical Care Medicine, University of Pittsburgh Medical Center, Pittsburgh*
[3] *Department of Emergency Medicine, University of Pittsburgh Medical Center, Pittsburgh*

Editors: Under Review for MIDL 2022

Abstract

We present the first known system to use robotic ultrasound to accurately map multiple anatomic structures in the femoral region in 3D and choose a point for vascular needle insertion that is suitable for catheter placement. The maps are presented as three-dimensional point clouds with points labeled categorically, e.g., veins, arteries, ligaments, etc. The final point for insertion is presented on a 3D map. We use a multi-class, multi-instance Bayesian 3D Convolutional Neural Network (CNN) to segment and identify the anatomic structures from 2D time series ultrasound data. The 2D segmented slices are then temporally stacked and synced with the kinematics of the robot maneuvering the ultrasound probe to create a 3D point cloud. This 3D point cloud is analyzed based on heuristics from physicians, to determine an ideal point to puncture with the needle, i.e., solve the needle-insertion planning problem. In particular, we determine the desired insertion point in either the common femoral artery or vein. We achieved a Jaccard score $J = \mathbf{0.834}$ for vessel segmentation and were able to determine safe insertion points in 46 out of 49 trials. Our system requires minimal human intervention and is designed to be robust to changes in ultrasound imaging settings and subject anatomy.

Keywords: Multi-class Segmentation, 3D Reconstruction, Needle Insertion, POCUS AI.

1. Introduction

Percutaneous needle-puncture procedures are used for a wide variety of anatomical targets within the body and are associated with performing safe and minimally-invasive surgeries. Common applications include central vascular access for resuscitation, arterial pressure

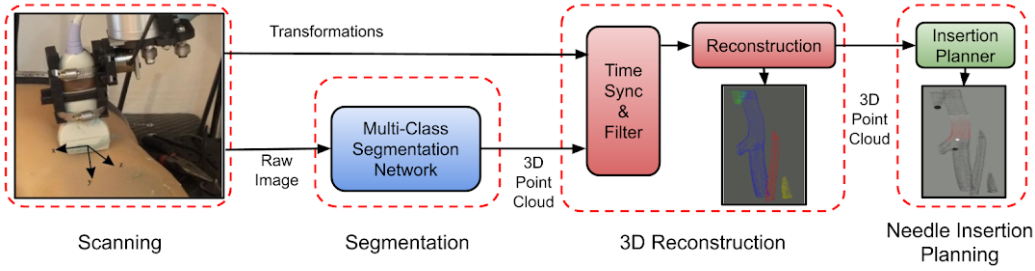

Figure 1: Flow diagram of the overall pipeline. The robot system collects ultrasound images from scanning, sends them to the deep learning model for segmentation, retrieves the segmentation coordinates from the generated 3D point cloud, and then outputs suggested locations for needle insertion based on safety standards.

monitoring, dialysis catheter placement as well as rarer, more invasive, endovascular interventions, Extracorporeal Membrane Oxygenation (ECMO), and Resuscitative Endovascular Balloon Occlusion (REBOA) (Tsui et al., 2008; Rao et al., 2008). In many of these procedures, placement of a needle in the appropriate location is essential to a positive outcome.

Literature on endovascular intervention supports that percutaneous femoral arterial access is associated with serious complications (Criado et al., 1998). Especially with older patients, complications related to insertion, such as hematomas (2-8%) and pseudoaneurysms (1-2%), become more common with the growing number of procedures done in the femoral area (Dudeck et al., 2004). The risks of complications increase further both in high-tempo, stressful situations and with less experienced clinicians. Furthermore, inaccurate judgement of mental 3D models or ultrasound images often result in multiple punctures, taking more time in critical scenarios. Severe medical issues also arise as a result of needle insertion in other location sites, such as transradial artery and liver access (Snelling et al., 2018). Automated approaches using robots can reduce these risks significantly (Chen et al., 2020).

Portability is key in emergency medical scenarios. Ultrasound probes are small, cost-effective, and field-portable, unlike the machines required for imaging techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), or X-ray. As a result, the robot can be easily transported across locations for serving emergency medical purposes.

In this paper, we present a robotic ultrasound imaging pipeline for vascular needle insertion in the femoral region as a major step towards real-world deployment during medical emergencies, outlined in Figure 1. The femoral region is preferred as it allows for rapid administration of medications critical in emergency situations (Castro et al., 2021). A sample set-up and result is shown in Figure 5 in Appendix B. For further elucidation of the algorithm's results, we also generate 3D heatmap visualizations depicting the needle insertion safety levels, along with uncertainty-based pruning of noisy segmentations. The robot is able to collect ultrasound images on both smooth and curved surfaces while being able to segment arteries, veins, ligaments, and nerves simultaneously. This entire pipeline can be executed with minimal human intervention, decreasing the expertise required to safely perform a variety of life-saving transcutaneous interventions (England et al., 2020). As a result, we introduce the following contributions:

1. A novel pipeline for standardizing an optimally safe location for vascular femoral-region needle insertion by using a 3D visualization generated from deep learning-based multi-class segmentations
2. A Bayesian framework for performing multi-class imaging, and 3D anatomic visualization all within a Bayesian framework

2. RELATED WORK

Currently, a wide variety of literature discusses robotic systems for needle insertion tasks (Cilingiroglu et al., 2011; Jayender et al., 2008; Smistad et al., 2017; Vogt et al., 2016; Oh et al., 2019; Castro et al., 2021; Chen et al., 2020). We can broadly categorize the approaches by the imaging modality and anatomical landmark localization method. For the femoral region, most of the existing works use non-ultrasound based imaging modalities, such as flouroscopy and X-ray (Cilingiroglu et al., 2011; Jayender et al., 2008), both of which expose the patient to ionizing radiation. In constrast, ultrasound is safe for continuous imaging and is also more portable, making it a better choice for our intended use case of emergency scenarios. Despite its lightweight benefits, ultrasound does have more noise in the images compared to other modalities. We deal with these characteristics by training a deep learning model based on the 3D U-Net (Ronneberger et al., 2015; Çiçek et al., 2016), within a Bayesian framework (Kendall and Gal; Kwon et al.), from various augmented images - helping it to learn discriminative features from the images.

Other existing works to localize anatomical landmarks rely on conventional methods such as brute force searches and vessel radii measurements. The non-deep learning methods rely on either an offline initial setup for registration of an *a priori* anatomic model to the images, or a brute force search for the image processing algorithm (Jayender et al., 2008; Smistad et al., 2017; Vogt et al., 2016; Smistad and Lindseth, 2016; Hadjerci et al., 2014). (Hadjerci et al., 2014) does a single class segmentation for nerves. (Chen et al., 2020) employs deep learning-based methods for vessel segmentation in a semi-automatic robotic pipeline using two near infrared and ultrasound imaging, for phlebotomy in the arm.

Some existing approaches to determine optimal needle insertion locations use geometrical models (Vogt et al., 2016). Other algorithms either don't take into account the location of the inguinal ligament or require manual input to obtain anatomical landmarks (Smistad et al., 2017; Oh et al., 2019; Chen et al., 2020). In prior works that use geometric models, the optimal insertion site is determined using a single image rather than a 3D visualization or sequence of frames from the insertion area. For the femoral region, it is critical to have an anatomical model of the area that considers the physical relationship of the global anatomical landmarks (Castro et al., 2021). Considering the locations of the inguinal ligament and vessel bifurcation is essential to prevent complications like retroperitoneal hematoma and hemorrhage (Bangalore and Bhatt, 2011; Burzotta et al., 2019).

The success of the aforementioned approaches lend themselves to the advantages of automated robotic needle insertion. However, these prior methods have limitations resulting from the use of multiple imaging modalities, classical segmentation approaches, and global anatomical landmark choices. These limitations preclude their portable use for real-world emergencies. Our proposed pipeline overcomes some of these deficiencies by learning to segment all relevant anatomical landmarks in the femoral region in a way that could be more universal for other imaging settings or human anatomical variations.

3. METHODS

3.1. Experimental Setup

We use the Universal Robot UR3e model for ultrasound scanning. A Fukuda Denshi portable point-of-care ultrasound scanner (POCUS), with a 5-12 MHz 2D transducer is used for scanning. The data was gathered from a CAE Blue Phantom anthropomorphic gel model, *blue-gel*, the left and right sides of a CAE Blue Phantom lower torso ultrasound training model BPF1500-HP having different anatomical variations, *torso-left* and *torso-right*, and live pigs, *live-pig*. The IACUC-approved experiments on live pigs were done in a controlled lab setting under clinician supervision. The deep learning pipeline for multi-class segmentation was built using TensorFlow (Abadi et al., 2016) and Python. Our optimal needle insertion location prediction algorithm was implemented in Python, and Robot Operating System (ROS) (Quigley et al.) was used to combine all of the components.

3.2. Multi-Class Segmentation of Arteries, Veins, Ligaments, and Nerves

The deep learning model we use is a Bayesian formulation of the 3D U-Net encoder-decoder architecture, inspired by (Kendall and Gal; Çiçek et al., 2016). We use a sequence of 8 two-dimensional ultrasound images as input to the model, where the temporal aspect is treated as the third dimension. Due to memory limits, the encoder side of the network consists of four encoder blocks, with each block consisting of 3D convolution, batch normalization, and ReLU layers as described in (Çiçek et al., 2016). The decoder side of the network consists of the encoder-paired decoder blocks (Çiçek et al., 2016). We formulate a Bayesian version (Kendall and Gal) of this 3D U-Net by placing a distribution over its weights with a single dropout layer at the output of each encoder and decoder block, which we empirically found to produce the best results. The model then consists of two outputs, one for the predictive mean, $\hat{\mu}$, and another for the predictive variance, $\hat{\sigma}^2$, as represented in (Kendall and Gal):

$$[\hat{\mu}, \hat{\sigma}^2] = f^{\hat{W}}(x) \tag{1}$$

where f is the Bayesian 3D U-Net, in this case, parameterised by model weights $\hat{W}$. The model is trained using the stochastic cross entropy loss formulated in (Kendall and Gal) that accounts for aleatoric uncertainty in the data. Epistemic uncertainty maps, which represent the model uncertainty, are obtained using test-time stochastic forward passes, also referred to as Monte Carlo dropout (Çiçek et al., 2016):

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{\mu}_t - \bar{\mu})^{\otimes 2} \tag{2}$$

where T is the total number of Monte Carlo samples and $\bar{\mu} = \sum_{t=1}^{T}\frac{\hat{\mu}_t}{T}$. Empirical trials and (Kendall and Gal) both illustrated that computing the logits variance output, $\hat{\sigma^2}$ was necessary, in spite of not actively using $\hat{\sigma^2}$ for computing the aleatoric uncertainty (which represents statistical uncertainty inherent in the data). We found that epistemic uncertainty overcompensates if the aleatoric uncertainty is not accounted for separately and we obtain poor performance.

To further account for variability in ultrasound imaging, data was augmented with rotations, translations, flipping, mirroring, zooming (in and out), filtering, and blurring prior to model training.

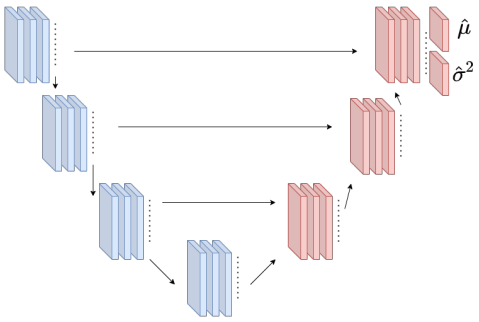

Figure 2: Diagram of our Bayesian 3D U-Net model.

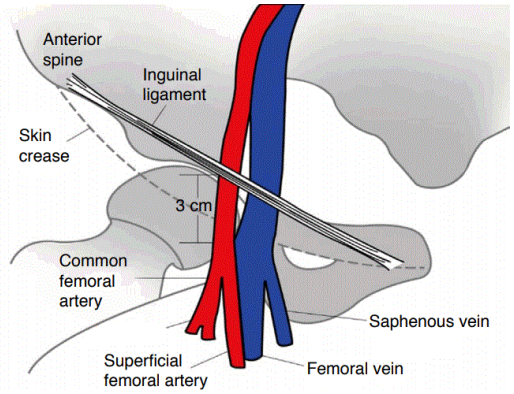

Figure 3: Diagram of ideal femoral arterial puncture site [15].

For training of the model, we used the Adam optimizer (Kingma and Ba, 2014), a learning rate value of 1e-4, which is dropped to 5e-5 after 30 epochs, a sequence length of 8 frames, and a batch size of 8. Our image resolution is 256x256. For obtaining the epistemic uncertainty maps, we use $T = 2$ Monte Carlo samples as it gave the best results. The model is shown in Figure 2.

3.3. 3D Visualization of Multi-Class Segmentation

We generate the 3D point cloud visualizations from the segmentation and uncertainty maps. The false-positive segmentations are filtered by: (1) calculating the average uncertainty values, υ_i, within every segmentation contour, (2) filtering υ_i by class, c, and calculating uncertainty thresholds, τ_c, with:

$$\tau_c = \hat{\upsilon_c} + \hat{\sigma_c} * \delta \tag{3}$$

where $\hat{\upsilon_c}$ and $\hat{\sigma_c}$ are the average and standard deviation, respectively, of υ_i taken for class c, and δ is a manually tuned parameter representing the number of standard deviations away from the mean to filter out. In practice, we found that the PERT statistical distribution (Clark, 1962) provided the best approximation to the uncertainty values υ_i.

The coordinates of the segmented vessels with respect to the robot base are calculated as follows:

$$^{ro}P_{im/px} = {}^{ro}T_{tr} * {}^{tr}T_{im/mm} * {}^{im/mm}S_{im/px} * p_{im/px} \tag{4}$$

where $p_{im/px}$ is the segmented region in the image, ${}^{im/mm}S_{im/px}$ is the scaling factor from pixels to millimeters (mm), ${}^{tr}T_{im/mm}$ is the transformation from mm units into the ultrasound transducer's frame, and ${}^{ro}T_{tr}$ is the transformation from the ultrasound transducer's frame to robot's fixed base frame. ${}^{tr}T_{im/mm}$ is obtained from the manual calibration procedure described in (Mercier, 2004) and ${}^{ro}T_{tr}$ is obtained from the *tf* ROS package.

3.4. Optimal Needle Insertion Location Prediction Algorithm

The femoral artery is our target vessel, but our pipeline can be easily extended to other vessels. A diagram of the anatomy is displayed in Figure 3. The ideal site for femoral arterial puncture is generally accepted to be over the femoral head, below the inguinal ligament, and above the femoral arterial bifurcation (Cilingiroglu et al., 2011), (Rupp et al., 1993).

Specifically, recent literature points towards safest insertion being at points lying within 58-92% of the way down from the top of the femoral head (femoral head divided cranio-caudally). In this region, the inguinal ligament and arterial bifurcation are typically 33% and 100% from the top of the femoral head, respectively (Gopalakrishnan et al., 2017). Even in cases where doctors are aware of the arterial bifurcation point through ultrasound imaging, they tend to aim at least 1 cm cranially from the bifurcation (Oğuzkurt et al., 2012). These practices are reflected in our algorithm. To further account for noise in the segmentation outputs, we filter out segmentation predictions for class c which have pixel-areas smaller than ϕ_c pixels. After gauging average pixel-areas of segmented vessels, we used 100, 300, and 1000 for ϕ_{artery}, ϕ_{vein}, and $\phi_{ligament}$, respectively, in the algorithm below:
(1) Detection of the femoral arterial bifurcation point: We implement this by checking for a gap, at least, of size g between the two detected contours from the artery class. We then assume the location with the smallest gap as the point of bifurcation and refer to this as point α. To account for noise in the segmentation results, we check that at least γ% of the image frames caudal to that point also contain at least 2 contours. From multiple experiments, we determined the values of g and γ to be 3 and 95% respectively.
(2) Detect the caudal end of the inguinal ligament: *If a ligament was detected* in the scan, we determine the closest point on the ligament to point α. *If the ligament was not scanned/detected*, we account for 2 scenarios: (a) the femoral artery follows a straight path, and (b) there exists a gradual curve in the femoral artery, referring to where the artery is crossing under the ligament. For (a), we currently assume the ligament's location to be immediately off the cranial edge of the scan. For (b), we iterate over each arterial contour at index i and calculate the angles between vectors u and v, with opposing endpoints at $i-k$ and $i+k$, respectively, using the following: $\cos\theta = \langle u_k, v_k\rangle / |u_k||v_k|$. An illustration is depicted in Figure 7 in Appendix C. We then assume the location at index i with the smallest angle θ as the ligament landmark, which we will refer to as point λ. We account for noise similar to the previous step, except by checking for a count of 1.
(3) Determine a safe region in between the two anatomical landmarks: We do this by shifting α and λ towards each other by δ_α% and δ_λ%, respectively. We denote these shifted safe boundaries as α_s and λ_s, respectively. We do this (1) to account for noise in the deep learning segmentation outputs and (2) to incorporate common medical practices as mentioned above. We used 15 for the value of δ_α and δ_λ, accounting for the 1 cm minimum distance to the arterial bifurcation given an average common femoral artery segment length of ∼7 cm (Garrett et al., 2005).
(4) Calculate the overlap percentage between the femoral vein and artery at all points: We do this by calculating the percentage of overlapping pixels when viewing the contours from a posterior angle. We refer to this as V_o.
(5) Compute scores and determine optimal insertion location: We use the following relations:

$$P_h(\zeta) = \begin{cases} \frac{\sigma_\alpha}{1+\|\zeta-\alpha\|_2} + \frac{\sigma_\lambda}{1+\|\zeta-\lambda\|_2}, & \text{if } \zeta \in [\lambda_s, \alpha_s] \\ \infty, & \text{otherwise} \end{cases}$$
$$T_s = P_h + V_o \tag{5}$$

where P_h is the *Proximity Hazard Score* to account for distance from α and λ, ζ is the 3D coordinate for the center of an arterial segmentation contour, σ_α and σ_λ are values for weighing the importance of sufficient distance from α and λ, respectively. We reflect the aforementioned insertion percentiles commonly used in practice with σ_α and σ_λ. Arterial

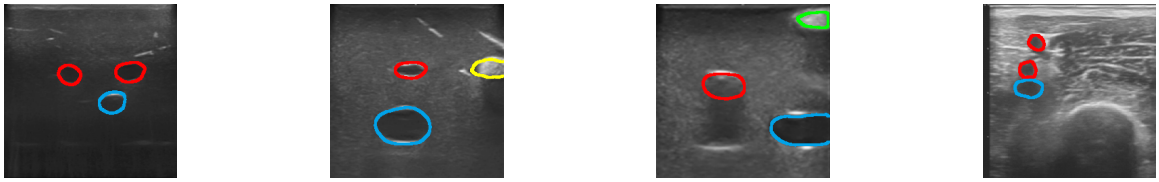

Figure 4: Segmentation results on images collected with different imaging settings and anatomy. Left to right: *blue-gel*, *torso-left*, *torso-right*, *live-pig*.

contours not within the safe region or of an area smaller than ϕ_{artery} are assigned a maximum score. The *Total Site Score*, T_s, is then obtained by adding V_o and P_h together, taking into account overflow and underflow. For the final insertion location, we select the artery corresponding to the lowest value of T_s. If there are multiple such values, we choose the largest-sized artery cross-section.

To enhance clinical viability and explainability, we generate a second 3D visualization illustrating a heatmap of T_s. We do this using the same steps described in the previous section, except with the following color encoding scheme: non-artery/vein structures and regions with values of ∞ are colored with RGB value (128,128,128), i.e., grey, and arteries are shaded with RGB values of (255,η,η) where $\eta = min(T_s - min(T_s) * 255, 255)$. Therefore, regions with higher values of T_s appear white, whereas lower values appear bright red.

4. ANALYSIS AND RESULTS

4.1. Multi-Class Segmentation

To validate the robustness of our multi-class segmentation network, we evaluate its performance on *blue-gel*, *torso-left*, and *torso-right* phantoms, and on *live-pig* datasets, all of which (except for *live-pig* which was just with im_a) are compared across 3 varying imaging settings: (im_a) depth of 5 cm and gain value of 15 units, (im_b) depth of 10 cm and gain value of 15 units, and (im_c) a depth of 10 cm with a gain value of 10 units. Each of the phantom datasets contained a train/valid/test split of 320/128/256 images, whereas the live-pig dataset contained a split of 1360/240/320 images, sourced from 8 different pigs. The images were split into sequential groups of 8 frames each. After applying data augmentation to the training images, as in (Mathai et al., 2019), the cardinality of each set grew by a factor of ~ 20. The outputs from the final layers were converted to color-coded masks using a threshold value of 0.50. Erosion is used to convert the dense segmentation mask into just the border of the circle. Image samples are shown in Figure 5 in Appendix B.

For baseline comparisons, we evaluated the results against those of a vanilla 3D U-Net, which we refer to as $3DU$, (Çiçek et al., 2016) and a different variation of the Bayesian 3D U-Net similar to that in (Kwon et al.), which we refer to as $B3DU_k$. The metrics we used are region similarity, J, and contour accuracy, F, (Clark, 1962). We evaluate the spatial region and contour segmentation performance of estimated segmentation, S and ground-truth mask, G, using region similarity and contour accuracy. The multi-class segmentation model performances are described in Table 1. As can be seen, *our model obtains the most accurate boundary contours for the blue-gel and torso-right phantoms.* Across all 4 anatomies, our model is the most consistent. Sample segmentation outputs are shown in Figure 4.

4.2. Optimal Needle Insertion Location Prediction

To test for robustness, 9 trials of 3D visualization and the optimal needle insertion location prediction algorithm were performed across *blue-gel*, *torso-left*, and *torso-right* phantoms, with one scan from each imaging setting im_a, im_b, and im_c. We also performed 8 trials for different sequences of *live-pig* scans. We assumed a venous insertion in *torso-left* and *torso-right* due to anatomical differences in the phantoms, whereas *blue-gel* and *live-pig* were arterial insertions. For *live-pig* data, we tune k, the gap of arterial contours for the endpoints of $\overrightarrow{u}$ and $\overrightarrow{v}$ (Figure 7 in Appendix C), to be within 1-5.

To evaluate our proposed insertion location prediction algorithm, we asked 3 doctors to judge the accuracy of the results. We initially only showed the doctors the unmarked 3D visualizations, (without insertion points depicted), and asked them mark safe regions for arterial/venous insertion along with the respective vessel bifurcation and ligament points. We then presented to them the 3D visualizations and heatmaps, with safe insertion points, bifurcation, and ligament points depicted, and asked them to confirm the correctness. On *blue-gel*, *torso-left*, and *torso-right*, our algorithm determined the anatomical landmarks and insertion points correctly **9/9** times. On *live-pig*, our algorithm correctly determined the ligament **8/8** times, the arterial bifurcation **7/8** times, and a safe insertion point **6/8** times. The 2 sequences without a safe insertion point were instead deemed slightly too close to the arterial bifurcation. Overall, our algorithm detected the proper arterial bifurcation and ligament landmarks **31/32** times, while detecting a safe insertion point **15/17** times. For additional validation, we also displayed plots illustrating the range and general distribution of proximity hazard scores, overlap scores, and total site scores. Computationally, just the insertion planning algorithm takes ≈ 0.0500 seconds to complete, averaged over 10 runs. Figures 5, 6, 7, and 8 in Appendices B and C illustrate our results.

5. CONCLUSION AND FUTURE WORK

We present a novel pipeline (Figure 1) for predicting safe needle insertion points in femoral vessels based on 3D visualizations built using multi-class segmentation in a Bayesian framework. In our model, arteries, veins, and ligaments can be accurately segmented and used as anatomical landmarks for guiding needle insertion. For clinicians, this frees up the cognitive burden of determining them manually. Our 3D visualization produces explainable visual results. Our validation across several anatomical and scanning variations demonstrates that our pipeline is generalizable. We obtained safe insertion locations 93% of the time.

In the future, we will extend the current segmentation framework to the longitudinal view of the ultrasound scan. We also want to test our pipeline on more data from diverse anatomies.

Acknowledgments

This present work was sponsored in part by US Army Medical contracts W81XWH-19-C0083 and W81XWH-19-C0101. We are pursuing intellectual-property protection. Galeotti is an advisor to Activ Surgical, Inc. and is a founder of the startup Elio AI, Inc.

References

Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: A system for large-scale machine learning. *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016*, pages 265–283, 2016. doi: 10.5555/3026877.3026899.

Sripal Bangalore and Deepak L. Bhatt. Femoral arterial access and closure. *Circulation*, 124(5), 8 2011. ISSN 00097322. doi: 10.1161/CIRCULATIONAHA.111.032235. URL https://www.ahajournals.org/doi/abs/10.1161/circulationaha.111.032235.

Francesco Burzotta, Osama Shoeib, Cristina Aurigemma, and Carlo Trani. Angio-Guidewire-Ultrasound (AGU) Guidance for Femoral Access in Procedures Requiring Large Sheaths. *The Journal of invasive cardiology*, 31, 12 2019.

Danny Castro, LeeAnne M. Martin Lee, and Beenish S. Bhutta. Femoral Vein Central Venous Access. *StatPearls*, 8 2021. URL https://www.ncbi.nlm.nih.gov/books/NBK459255/.

Alvin I. Chen, Max L. Balter, Timothy J. Maguire, and Martin L. Yarmush. Deep learning robotic guidance for autonomous vascular access. *Nature Machine Intelligence 2020 2:2*, 2(2):104–115, 2 2020. ISSN 2522-5839. doi: 10.1038/s42256-020-0148-7. URL https://www.nature.com/articles/s42256-020-0148-7.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *undefined*, 9901 LNCS:424–432, 2016. ISSN 16113349. doi: 10.1007/978-3-319-46723-8{_}49. URL http://lmb.informatik.uni-freiburg.de/resources/opensource/unet.en.html.

Mehmet Cilingiroglu, Ted Feldman, Michael H. Salinger, Justin Levisay, and Zoltan G. Turi. Fluoroscopically-guided micropuncture femoral artery access for large-caliber sheath insertion. *The Journal of Invasive Cardiology*, 23(4):157–161, 4 2011. ISSN 1042-3931. URL https://europepmc.org/article/med/21474850.

Charles E. Clark. Letter to the Editor—The PERT Model for the Distribution of an Activity Time. *https://doi.org/10.1287/opre.10.3.405*, 10(3):405–406, 6 1962. ISSN 0030-364X. doi: 10.1287/OPRE.10.3.405. URL https://pubsonline.informs.org/doi/abs/10.1287/opre.10.3.405.

Frank J. Criado, Mordechai Twena, Maria Halsted, and Omran Abul-Khoudoud. Percutaneous femoral puncture for endovascular treatment of occlusive arterial lesions. *American journal of surgery*, 176(2):119–121, 8 1998. ISSN 0002-9610. doi: 10.1016/S0002-9610(98)00162-7. URL https://pubmed.ncbi.nlm.nih.gov/9737614/.

Oliver Dudeck, Ulf Teichgraeber, Petr Podrabsky, Enrique Lopez Haenninen, Reingard Soerensen, and Jens Ricke. A randomized trial assessing the value of ultrasound-guided puncture of the femoral artery for interventional investigations. *The international journal of cardiovascular imaging*, 20(5):363–368, 10 2004. ISSN 1569-5794. doi: 10.1023/B:CAIM.0000041949.59255.3F. URL https://pubmed.ncbi.nlm.nih.gov/15765858/.

Elizabeth C. England, Charlotte R. Spear, Dih Dih Huang, Jordan Weinberg, James N. Bogert, Thomas Gillespie, and James Mankin. REBOA as a rescue strategy for catastrophic vascular injury during robotic surgery. *Journal of robotic surgery*, 14(3):473–477, 6 2020. ISSN 1863-2491. doi: 10.1007/S11701-019-01011-3. URL https://pubmed.ncbi.nlm.nih.gov/31463880/.

Paul D. Garrett, Robert E. Eckart, Terry D. Bauch, Christopher M. Thompson, and Karl C. Stajduhar. Fluoroscopic localization of the femoral head as a landmark for common femoral artery cannulation. *Catheterization and cardiovascular interventions : official journal of the Society for Cardiac Angiography & Interventions*, 65(2):205–207, 6 2005. ISSN 1522-1946. doi: 10.1002/CCD.20373. URL https://pubmed.ncbi.nlm.nih.gov/15900552/.

Prabhakaran Gopalakrishnan, Pradeep Manoharan, Chander Shekhar, Sharan Rufus Ponniah, Muhammad Chaudhry, Pradeepto Ghosh, Digant Bhatt, Ira Friedlander, and Nagapradeep Nagajothi. MID FEMORAL HEAD IS NOT THE IDEAL FLUROSCOPIC LANDMARK FOR COMMON FEMORAL ARTERY PUNCTURE. *Journal of the American College of Cardiology*, 69(11):1158, 3 2017. ISSN 07351097. doi: 10.1016/S0735-1097(17)34547-3.

Oussama Hadjerci, Adel Hafiane, Pascal Makris, Donatello Conte, Pierre Vieyres, and Alain Delbos. Nerve detection in ultrasound images using median gabor binary pattern. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8815:132–140, 2014. ISSN 16113349. doi: 10.1007/978-3-319-11755-3{_}15. URL https://www.researchgate.net/publication/269102105_Nerve_Detection_in_Ultrasound_Images_Using_Median_Gabor_Binary_Pattern.

Jagadeesan Jayender, Mahdi Azizian, and Rajini V. Patel. Autonomous Image-Guided Robot-Assisted Active Catheter Insertion. *IEEE Transactions on Robotics*, 24(4):858–871, 8 2008. ISSN 15523098. doi: 10.1109/TRO.2008.2001353. URL https://dl.acm.org/doi/abs/10.1109/TRO.2008.2001353.

Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

Diederik P. Kingma and Jimmy Lei Ba. Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 12 2014. URL https://arxiv.org/abs/1412.6980v9.

Yongchan Kwon, Joong-Ho Won, Beom Joon Kim, and Myunghee Cho Paik. Uncertainty quantification using Bayesian neural networks in classification: Application to ischemic stroke lesion segmentation.

Tejas Sudharshan Mathai, Vijay Gorantla, and John Galeotti. Segmentation of Vessels in Ultra High Frequency Ultrasound Sequences Using Contextual Memory. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11765 LNCS:173–181, 2019. ISSN 16113349. doi: 10.1007/978-3-030-32245-8{_}20.

Laurence Mercier. Review of ultrasound probe calibration techniques for 3D ultrasound. 12 2004.

Ting Ting Oh, Mohammad Ikhsan, Kok Kiong Tan, Sultana Rehena, Nian Lin Reena Han, Alex Tiong Heng Sia, and Ban Leong Sng. A novel approach to neuraxial anesthesia: Application of an automated ultrasound spinal landmark identification. *BMC Anesthesiology*, 19(1):1–8, 4 2019. ISSN 14712253. doi: 10.1186/S12871-019-0726-6/FIGURES/4. URL https://bmcanesthesiol.biomedcentral.com/articles/10.1186/s12871-019-0726-6.

Levent Oğuzkurt, Kamil Gürel, Evren Eker, Serkan Gür, Uğur Özkan, and Öner Gülcan. Ultrasound-guided puncture of the femoral artery for total percutaneous aortic aneurysm repair. *Diagnostic and interventional radiology (Ankara, Turkey)*, 18(1):92–95, 1 2012. ISSN 1305-3612. doi: 10.4261/1305-3825.DIR.4061-10.1. URL https://pubmed.ncbi.nlm.nih.gov/21674453/.

Morgan Quigley, Brian Gerkey, Ken Conley, Josh Faust, Tully Foote, Jeremy Leibs, Eric Berger, Rob Wheeler, and Andrew Ng. ROS: an open-source Robot Operating System. URL http://stair.stanford.edu.

Sunil V. Rao, Fang Shu Ou, Tracy Y. Wang, Matthew T. Roe, Ralph Brindis, John S. Rumsfeld, and Eric D. Peterson. Trends in the Prevalence and Outcomes of Radial and Femoral Approaches to Percutaneous Coronary Intervention. A Report From the National Cardiovascular Data Registry. *JACC: Cardiovascular Interventions*, 1(4):379–386, 8 2008. ISSN 19368798. doi: 10.1016/J.JCIN.2008.05.007.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351:234–241, 2015. ISSN 16113349. doi: 10.1007/978-3-319-24574-4{_}28. URL https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28.

Stephen B. Rupp, Robert L. Vogelzang, Albert A. Nemcek, and Margaret M. Yungbluth. Relationship of the inguinal ligament to pelvic radiographic landmarks: anatomic correlation and its role in femoral arteriography. *Journal of vascular and interventional radiology : JVIR*, 4(3):409–413, 1993. ISSN 1051-0443. doi: 10.1016/S1051-0443(93)71889-8. URL https://pubmed.ncbi.nlm.nih.gov/8513216/.

Erik Smistad and Frank Lindseth. Real-Time Automatic Artery Segmentation, Reconstruction and Registration for Ultrasound-Guided Regional Anaesthesia of the Femoral Nerve. *IEEE transactions on medical imaging*, 35(3):752–761, 3 2016. ISSN 1558-254X. doi: 10.1109/TMI.2015.2494160. URL https://pubmed.ncbi.nlm.nih.gov/26513782/.

Erik Smistad, Daniel Høyer Iversen, Linda Leidig, Janne Beate Lervik Bakeng, Kaj Fredrik Johansen, and Frank Lindseth. Automatic Segmentation and Probe Guidance for Real-Time Assistance of Ultrasound-Guided Femoral Nerve Blocks. *Ultrasound in medicine & biology*, 43(1):218–226, 1 2017. ISSN 1879-291X. doi: 10.1016/J.ULTRASMEDBIO.2016.08.036. URL https://pubmed.ncbi.nlm.nih.gov/27727021/.

Brian M. Snelling, Samir Sur, Sumedh Subodh Shah, Megan M. Marlow, Mauricio G. Cohen, and Eric C. Peterson. Transradial access: Lessons learned from cardiology. *Journal of NeuroInterventional Surgery*, 10(5):493–498, 1 2018. ISSN 1759-8478. doi: 10.1136/NEURINTSURG-2017-013295. URL https://miami.pure.elsevier.com/en/publications/transradial-access-lessons-learned-from-cardiology.

Janet Y. Tsui, Adam B. Collins, Douglas W. White, Jasmine Lai, and Jeffrey A. Tabas. Videos in clinical medicine. Placement of a femoral venous catheter. *The New England journal of medicine*, 358(26):e30, 6 2008. ISSN 1533-4406. doi: 10.1056/NEJMVCM0801006. URL https://pubmed.ncbi.nlm.nih.gov/18579807/.

Mark Vogt, Dennis J. Van Gerwen, John J. Van Den Dobbelsteen, and Martin Hagenaars. Optimal point of insertion of the needle in neuraxial blockade using a midline approach: study in a geometrical model. *Local and regional anesthesia*, 9:39–44, 8 2016. ISSN 1178-7112. doi: 10.2147/LRA.S111569. URL https://pubmed.ncbi.nlm.nih.gov/27570462/.

Appendix A. Tables

Table 1: Evaluation with region similarity, computed as the Jaccard score (J) and contour accuracy, computed as the F measure over precision and recall (F) with the obtained segmentations and the ground truth annotations. Arrows indicate optimal direction for avg ± std across im_a, im_b, im_c (im_a for *live-pig*).

.

Model	$J \uparrow$	$F \uparrow$
blue-gel		
$3DU$	.775 ± .035	.291 ± .050
$B3DU_k$	.785 ± .015	.316 ± .064
$Ours$	**.834 ± .008**	**.516 ± .008**
torso-left		
$3DU$	.663 ± .014	.442 ± .025
$B3DU_k$	**.793 ± .107**	.656 ± .140
$Ours$	.788 ± .114	**.662 ± .140**
torso-right		
$3DU$	.706 ± .076	.402 ± .007
$B3DU_k$	.807 ± .041	.571 ± .032
$Ours$	**.816 ± .057**	**.657 ± .013**
live-pig		
$3DU$	**.84 ± .132**	.693 ± .071
$B3DU_k$	.795 ± .170	.667 ± .071
$Ours$	.828 ± .153	**.768 ± .054**

Appendix B. Segmentation Results

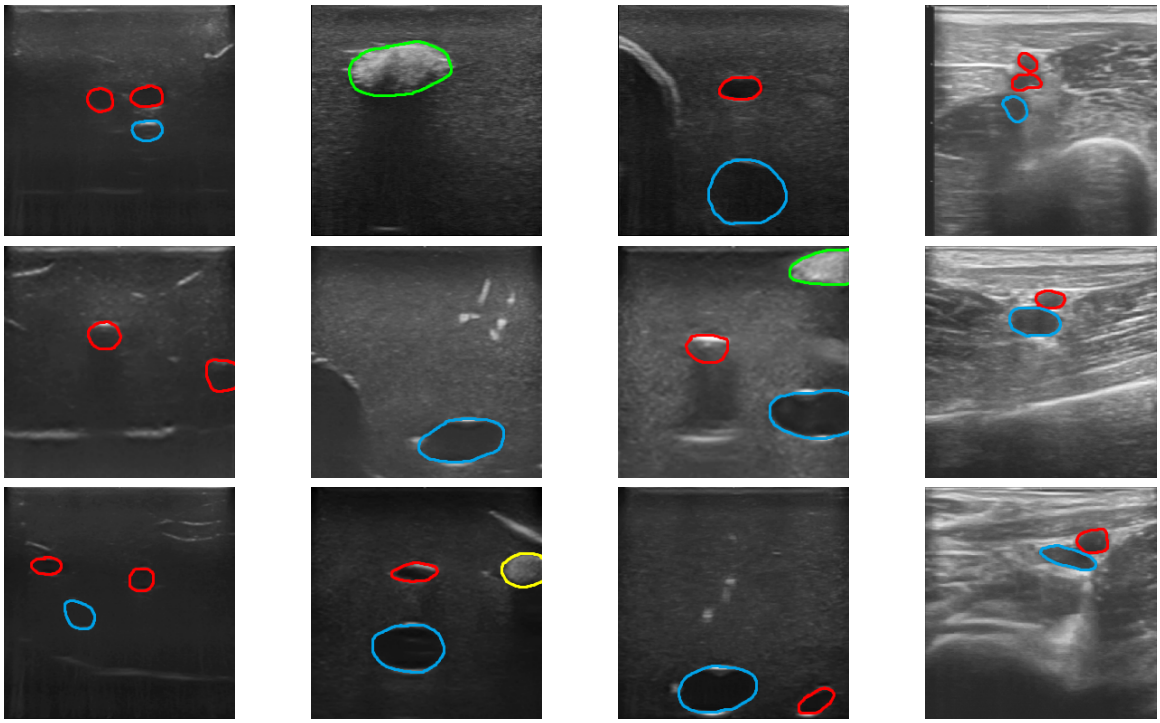

Figure 5: Examples of the variety in images we validated our methods on, with the labels overlaid. From left to right: *blue-gel*, *torso-left*, *torso-right*, *live-pig*. Each row shows the 3 different imaging settings with which we varied for each subject, except for *live-pig*. **Color Key**: Arteries - Red, Veins - Blue, Ligaments - Green, Nerves - Yellow.

Appendix C. Visualization of 3D Reconstructions

This section represents visualization of *blue gel*, *live pig* and *torso right*. The optimal insertion point is shown by the white dot, whereas the ligament and arterial bifurcation points are the grey dots.

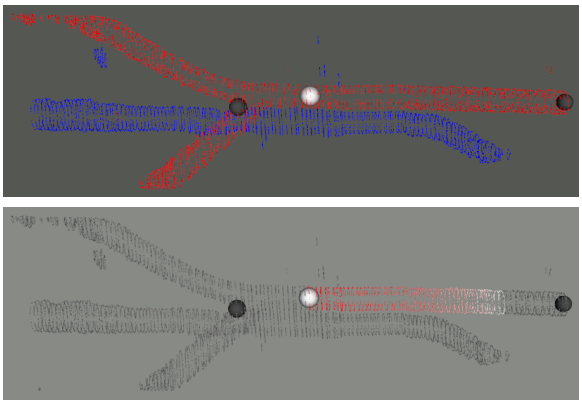

Figure 6: 3D visualization of *blue-gel*. The image on the bottom illustrates a heatmap of the *Total Site Scores* for the needle planning algorithm (unsafe to safe goes from gray to red).

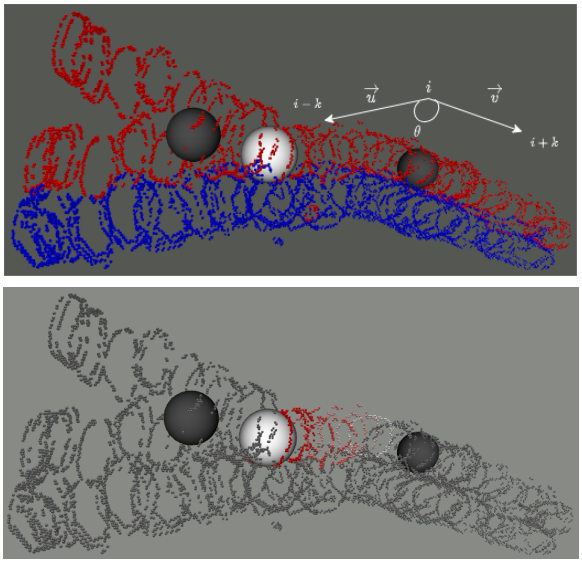

Figure 7: 3D visualization of *live-pig*. The top image shows the method we use for detecting the ligament. The image on the bottom illustrates a heatmap of the *Total Site Scores* for the needle planning algorithm (unsafe to safe goes from gray to red).

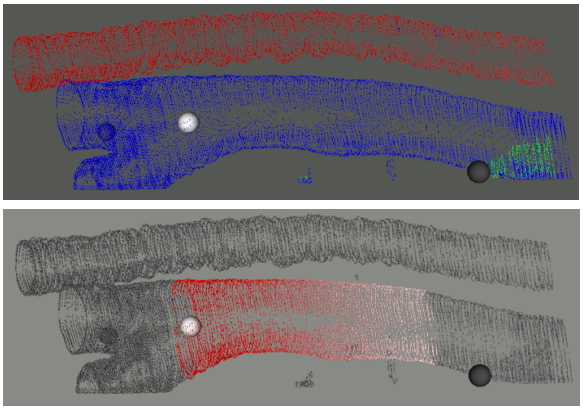

Figure 8: 3D visualization of *torso-right*. The image on the bottom illustrates a heatmap of the *Total Site Scores* for the needle planning algorithm (unsafe to safe goes from gray to red).