

# **Nerve Block Target Localization and Needle Guidance for Autonomous Robotic Ultrasound Guided Regional Anesthesia**

Thesis submitted in partial fulfillment  
of the requirements for the degree of

Master of Science  
in  
Computer Science and Engineering by Research

by

**ABHISHEK TYAGI**

20173067

`abhishek.tyagi@research.iiit.ac.in`



International Institute of Information Technology, Hyderabad

Hyderabad - 500 032, INDIA

November 2023

Copyright © ABHISHEK TYAGI, 2023  
All Rights Reserved

International Institute of Information Technology  
Hyderabad, India

**CERTIFICATE**

It is certified that the work contained in this thesis, titled “NERVE BLOCK TARGET LOCALIZATION AND NEEDLE GUIDANCE FOR AUTONOMOUS ROBOTIC ULTRASOUND GUIDED REGIONAL ANESTHESIA” by ABHISEK TYAGI, has been carried out under my supervision and is not submitted elsewhere for a degree.

---

Date

---

Adviser: Prof. JAYANTHI SIVASWAMY

To my family,  
who stood supportive of my decision to pursue higher studies.

## **Acknowledgments**

First and foremost, I would like to express my sincere gratitude to my research advisor Prof. Jayanthi Sivaswamy, for providing valuable guidance and support throughout my research journey. In spite of her busy schedule, she is always reachable and is genuinely interested towards the upliftment of her students. During the challenging pandemic times, she was connected with all students. Her teaching style and depth inspires motivation. As one of the distinguished researchers, she taught us to connect the traditional computer vision concepts with modern machine learning to understand the medical image field holistically. I will always be grateful to her and I hope I can stay true to the principles she has taught us.

I feel fortunate to be a part of the extraordinary group of medical image researchers at CVIT who are always eager to help each other. I have very fond memories and will always cherish the time spent with Naren, Alphin, Samruddhi, Lokesh, Pratyusha, Aniket and Abhinav.

I want to take this opportunity to thank my wife Neha and brother Abhay for always supporting me to undertake this research journey. I thank my parents and parents-in-law for encouragement and blessings. I am thankful to my daughter Ananya for kindling curiosity which is always helpful to a researcher. I owe a debt of gratitude to Aman, Bineet, Khushbu and Divyanshi for all the help they provided.

## Abstract

Ultrasound guided regional anesthesia (UGRA) involves approaching target nerves through a needle in real-time, enabling precise deposition of drug with increased success rates and fewer complications. Development of autonomous robotic systems capable of administering UGRA is desirable for remote settings and localities where anesthesiologists are unavailable. Real-time segmentation of nerves, needle tip localization and needle trajectory extrapolation are required for developing such a system. In the first part of this thesis, we developed models to localize nerves in the ultrasound domain using a large dataset. Our prospective study enrolled 227 subjects who were systematically scanned for brachial plexus nerves in various settings using three different ultrasound machines to create a dataset of 227 unique videos. In total, 41,000 video frames were annotated by experienced anaesthesiologists using partial automation with object tracking and active contour algorithms. Four baseline neural network models were trained on the dataset and their performance was evaluated for object detection and segmentation tasks. Generalizability of the best suited model was then tested on the datasets constructed from separate ultrasound scanners with and without fine-tuning. The results demonstrate that deep learning models can be leveraged for real time segmentation of brachial plexus in neck ultrasonography videos with high accuracy and reliability. Using these nerve segmentation predictions, we define automated anesthesia needle targets by fitting an ellipse to the nerve contours. The second part of this thesis focuses on localization of the needles and development of a framework to guide the needles toward their targets. For the segmentation of the needle, a natural RGB pre-trained neural network is first fine-tuned on a large ultrasound dataset for domain transfer and then adapted for the needle using a small dataset. The segmented needle's trajectory angle is calculated using Radon transformation and the trajectory is extrapolated from the needle tip. The intersection of extrapolated trajectory with the needle target guides the needle navigation for drug delivery. The needle trajectory's average angle error was  $2^\circ$ , average error in trajectory's distance from center of the image was 10 pixels (2 mm) and the average error in needle tip was 19 pixels (3.8 mm) which is within acceptable range of 5 mm as per experienced anesthesiologists. The entire dataset has been released publicly for further study by the research community.

# Contents

<b>1. Introduction.....</b>	<b>11</b>
1.1 Peripheral Nerve Blocks and Ultrasound Guided Regional Anesthesia.....	11
1.2 Motivation.....	12
1.3 Brachial Plexus.....	12
1.4 Outline.....	15
<b>2. Automated Real Time Delineation of Brachial Plexus in Ultrasonography.....</b>	<b>16</b>
2.1 Introduction.....	16
2.2. Methods.....	19
2.2.1 Dataset curation.....	19
2.2.2 Data augmentation.....	22
2.2.3 Neural Network.....	23
2.2.4 Internal validation.....	25
2.2.5 Generalizability.....	25
2.2.6 Evaluation.....	25
2.3. Results.....	26
2.4. Discussion.....	29
<b>3. Needle Guidance for Autonomous Robotic Ultrasound Guided Regional Anesthesia.....</b>	<b>33</b>
3.1 Introduction.....	33
3.1.1 Ultrasound Challenges.....	33
3.1.2 Recent work.....	35
3.1.3 Contribution.....	35
3.2 Materials and Methods.....	36
3.2.1 Dataset and Annotation.....	36
3.2.2 Nerve block Target.....	36
3.2.3 Neural Network.....	37

3.2.4 Needle Trajectory.....	37
3.2.5 Evaluation.....	41
3.3 Results.....	41
3.4 Discussion and Conclusion.....	42
<b>4. Conclusions.....</b>	<b>44</b>
4.1 Future Work.....	45
Bibliography.....	47

## List of Figures

Figure	Page
1.1 Dissection showing the axillary blood-vessels and the brachial plexus of nerves.....	13
1.2 Brachial plexus nerves network.....	13
1.3 Interscalene block.....	14
1.4 Sagittal view of brachial plexus in ultrasound.....	15
2.1 Annotation process steps.....	21
2.2 Gamma intensity transformation.....	22
2.3 Synthetic acoustic shadow.....	23
2.4 Framework of neural network architecture.....	24
2.5 Precision-Recall Curves .....	28
3.1 Ellipse fitted on the contour of the nerve segmentation prediction.....	36
3.2 Needle target .....	37
3.3 Radon transform.....	38
3.4 Needle trajectory and sinogram.....	39
3.5. Needle trajectory guidance towards target.....	40
3.6 Needle trajectory evaluation parameters.....	41

## List of Tables

Table	Page
2.1 Data Set Patient Characteristics.....	20
2.2 Comparison of baseline models on Object Detection & Segmentation metrics.....	27
2.3 UNet++ performance on different machines.....	28

## **List of Supplementary Videos**

2.1 Algorithm segmentation (prediction mask) in blue compared to ground truth (red)

<https://youtu.be/iElBptnSOHA>

2.2 Model predictions differentiating SCBP and ISC with US probe movement along the neck. Ground Truth (red), SCBP mask (blue) and ISC mask (yellow)

<https://youtu.be/MMaDfMg-YVs>

2.3 Model interaction with needle insertion in Brachial Plexus

[https://youtu.be/Zn7\\_O4rgIhU](https://youtu.be/Zn7_O4rgIhU)

3.1 Needle Trajectory

<https://youtu.be/QUecX6iB3XI>

## Glossary

CNN	Convolutional Neural Network
DL	Deep Learning
IOU	Intersection Over Union
ISC	Interscalene Brachial Plexus
PNB	Peripheral nerve blocks
SCBP	Supraclavicular Brachial Plexus
SNR	Signal-to-noise ratio
UGRA	Ultrasound Guided Regional Anesthesia
US	Ultrasound

## *Chapter 1*

### **Introduction**

#### **1.1 Peripheral Nerve Blocks and Ultrasound Guided Regional Anesthesia**

A peripheral nerve block (PNB) is a shot of anesthesia medicine around a specific nerve or bundle of nerves to block pain. It can be used to numb a part of the body such as an arm or a leg that needs surgery for up to 24 hours [1]. Ultrasound guided regional anesthesia (UGRA) is a standard technique for administering peripheral nerve blocks. Ultrasound imaging allows doctors to see the anatomy of the area where they are injecting local anesthetic. This helps them to avoid damaging any important structures with the needle. Ultrasound also allows doctors to see the needle tip as it is being inserted, which helps to ensure that it is in the correct position. This is important because it reduces the risk of accidentally injecting the local anesthetic into a blood vessel or other unintended structure. In addition, ultrasound imaging allows doctors to see the spread of the local anesthetic as it is being injected. This helps to ensure that the anesthetic is distributed evenly, which provides better pain relief.

Some of the specific benefits of using ultrasound guidance for local anesthetic injections:

- Increased accuracy: Ultrasound guidance can help doctors to place the needle more accurately, which reduces the risk of complications.
- Improved safety: Ultrasound guidance can help doctors to avoid damaging important structures, such as blood vessels and nerves.
- Increased patient comfort: Ultrasound guidance can help to reduce the pain associated with local anesthetic injections.
- Improved success rate: Ultrasound guidance can help to increase the success rate of local anesthetic injections.

## 1.2 Motivation

The success of UGRA procedures in turn depends on accurate interpretation of sono-anatomy that can be challenging for medical trainees.

Development of autonomous robotic systems capable of administering UGRA is desirable for remote settings and localities where anesthesiologists are unavailable [47]. Real-time segmentation of nerves, needle tip localization and needle trajectory extrapolation are required for developing such a system.

Ultrasonography is generally considered a challenging modality for image segmentation because of poor contrast images, operator dependency, and the reduced SNR by granular image texture from speckle noise. In addition, for real-time segmentation, the hazy boundaries and speckle noise introduce additional time-complexity hurdles on image preprocessing. Needles occupy very few pixels in images which pose additional challenges as their appearance is often degraded by reverberation, comet tail, side-lobe, beam-width, or bayonet artifacts. If the needle does not lie parallel to the ultrasound beam's plane, it is visible only partially at the intersection with the plane. Also, the segmentation process may be disturbed by adjacent linear anatomical structures (e.g., borders of bones or fasciae pleura etc). The invasive nature of needles limits the possibility of acquiring large scale data required for deep learning methods.

In recent years, attempts have been made to identify individual nerves, surrounding structures and needles using object detection and segmentation neural networks [48]. However, most of these studies are limited by time complexity and small-sized training datasets that affects their accuracy and generalizability. Most of the automatic needle detection methods are based on phantoms or ex-vivo porcine and chicken tissues [48]. However, compared to phantoms, the human tissues are much more heterogeneous with many artifacts, underscoring the need to work with real in-vivo human data.

## 1.3 Brachial Plexus

Among the regional anesthesia targets for upper limb surgeries, two of the most commonly employed techniques include brachial plexus blockade at Interscalene (ISC) and supraclavicular space (SCBP).

The brachial plexus (Fig. 1.1) is a network of nerves that branches off the spinal cord (Fig. 1.2) and supplies the arm, shoulder, and hand. The interscalene block (Fig. 1.3) is a procedure that numbs the shoulder, arm, and clavicle by injecting local anesthetic into the nerves of the brachial plexus located in a

groove between the anterior and middle scalene muscles. The interscalene block is a commonly used procedure for providing anesthesia and postoperative analgesia for surgery to the shoulder and arm.

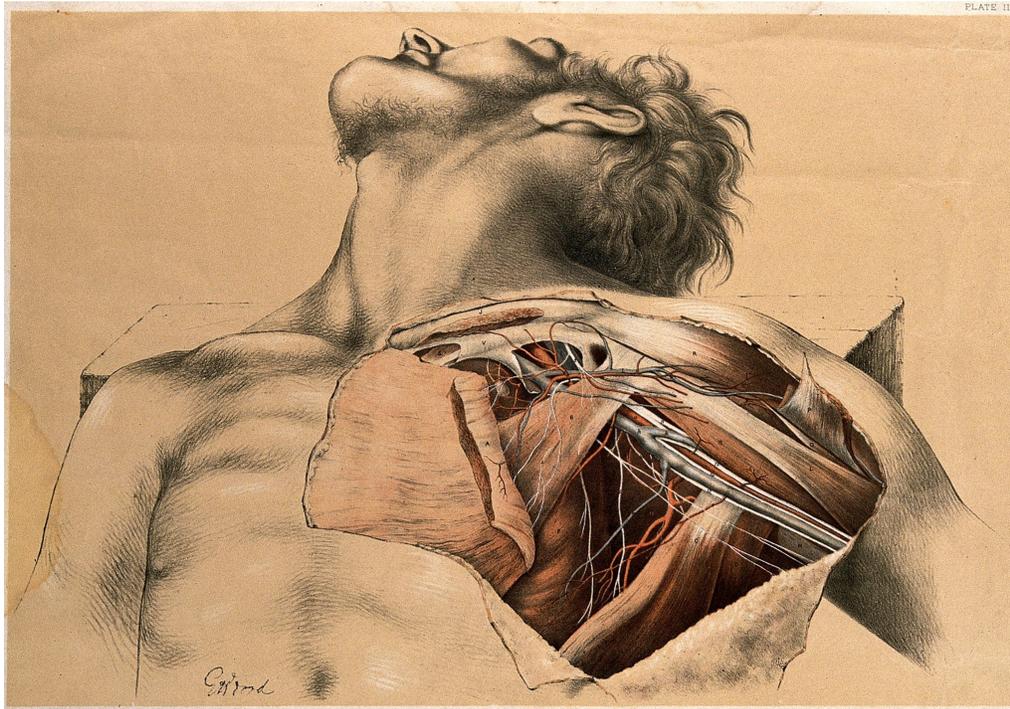


Fig 1.1. Dissection showing the axillary blood-vessels and the brachial plexus of nerves. Color lithograph by G.H. Ford, 1867. Credit: Wellcome Collection. Public Domain Mark

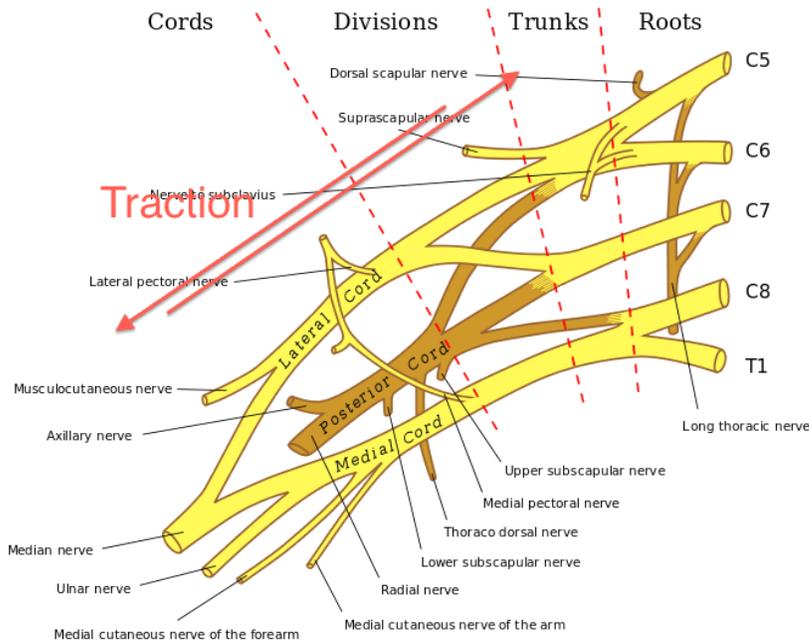


Fig 1.2 Brachial plexus nerves network by Donthatemebro, Wikimedia Commons CC BY-SA 3.0.

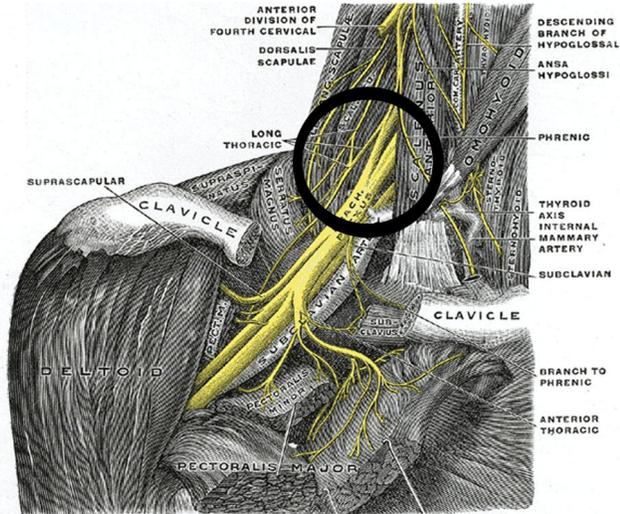


Fig 1.3 Left: the red line corresponds to the course of the subclavian artery, while the yellow line represents the brachial plexus and the "X" represents the site of entry of the needle when performing an interscalene block. Right: diagram of the course of the brachial plexus in relation to other important anatomic structures in the right side of the neck. Credit: David Shankbone (Foto), Grey's anatomy (2008). Wikimedia Commons under CC BY-SA 3.0.

The supraclavicular block is a regional anesthetic technique that numbs the arm and forearm from the lower humerus down to the hand. It is a single-injection block that is ideal for operations involving the upper extremity, excluding the shoulder. The supraclavicular block is performed by injecting local anesthetic into the brachial plexus at the level of the trunks, which are the three main branches of the plexus. This is the most compact level of the brachial plexus, so nerve block at this level has the greatest likelihood of blocking all of the branches of the plexus.

The supraclavicular block can be performed using surface landmarks or ultrasound guidance. The surface landmarks are the lateral border of the sternocleidomastoid muscle and the clavicle. The first rib generally represents the limit below which the needle must not be directed, as the pleural cavity and uppermost part of the lung are located at this level. Palpation or ultrasound visualization of the subclavian artery just above the clavicle provides a useful anatomic landmark for locating the brachial plexus.

The supraclavicular block is a safe and effective technique with few side effects. It is a good option for patients who need anesthesia and analgesia for surgery or other procedures involving the arm and forearm.

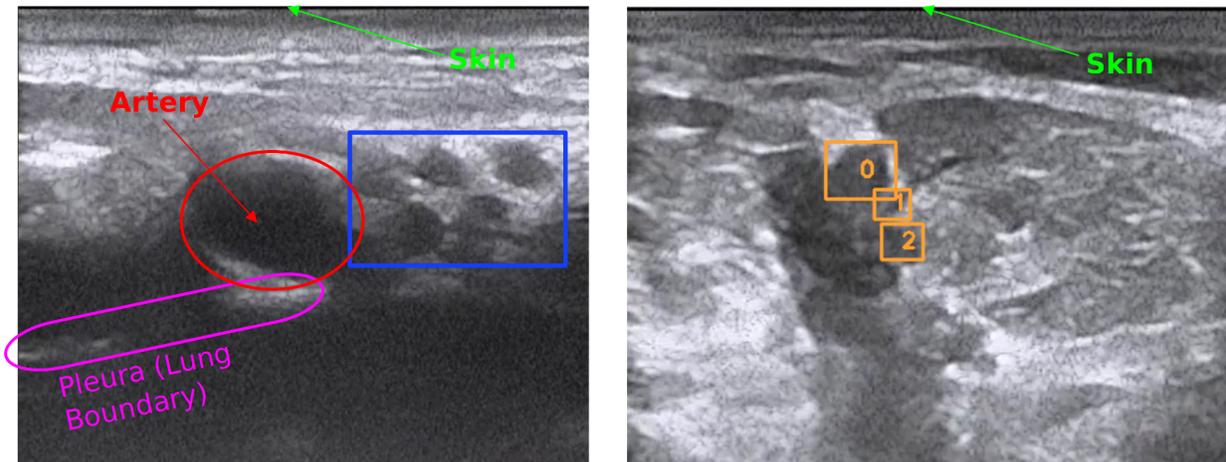


Fig. 1.4 Sagittal view of brachial plexus in ultrasound - Left: Supraclavicular, Right: Interscalene

## 1.4 Outline

In Chapter 2, an end-to-end deep learning approach for real time delineation of SCBP and ISC in neck ultrasound videos is presented.

In Chapter 3, we define automated anesthesia needle targets derived from CNN nerve segmentation predictions and design and implement automated needle segmentation, needle tip localization and needle trajectory extrapolation.

As part of the outcomes of this research, we also publicly released two datasets:

- 1) 227 brachial plexus ultrasound videos (41,000 frames) with segmentation annotations from experienced anesthesiologists
- 2) needle annotations of 6 ultrasound SCBP videos comprising of 1,061 frames

## *Chapter 2*

### **Automated Real Time Delineation of Brachial Plexus in Ultrasonography**

#### **2.1 Introduction**

Peripheral nerve blocks (PNB) are a critical component of multimodal approach for treating post-surgical pain. When compared to opioids alone, PNBs offer superior pain control and are associated with reduced perioperative opioid consumption, improved overall patient satisfaction and reduced hospital stay [1,2]. Over the past decade, use of ultrasound (US) has dramatically increased the scope of regional anesthesia. Ultrasound technology offers the ability to approach a target in real-time, enabling precise deposition of drug with increased success rates, reduced local anesthetic volume and fewer complications [3]. The success of US guided nerve blocks in turn, depends on accurate interpretation of sono-anatomy which can often be daunting for novice operators [4]. The classical approach of didactics, cadaveric demonstrations and 1:1 illustration are known to be cost-ineffective and time consuming. Simulation based techniques such as embedded electronic tutorials in US machines have consistently shown to be beneficial in enhancing sono-anatomical knowledge among trainees [3,5,6]. However, these techniques remain sporadic in practice with limitations in scope and can benefit from newer technologies.

Among the PNB targets to provide analgesia for upper limb surgeries, the two most commonly employed bread and butter techniques include brachial plexus blockade at interscalene and supraclavicular space [7]. These blocks target nerve roots and their divisions early on in their course, providing rapid and reliable anesthesia with minimal needling and patient discomfort. The interscalene brachial plexus (ISC) is blocked at the level of nerve roots exiting from the cervical spinal cord and provides analgesia to

shoulder region and upper arm whereas supraclavicular nerve block (SCBP) is performed inferiorly, lateral to the clavicular head of the sternocleidomastoid muscle and provides analgesia to the upper arm, forearm and hand [7,8]. When performing these blocks under real-time US guidance, the clinician looks for distinct sono-anatomical patterns, such as, cluster of hypoechoic nodules or bunch of grapes next to subclavian artery for SCBP and a vertical chain of hypoechoic nodules or a traffic light sign sandwiched between two scaleni muscles for ISC [9]. However, the aforementioned patterns only serve as general guide and can vary considerably depending upon individual patient's anatomy and US probe orientation. Moreover, these two sites are situated very close to each other (within a few centimeters) in the neck and transition seamlessly into one another making it challenging to distinguish for novice operators.

In recent years, convolutional neural networks (CNN) utilizing large sized training datasets have shown remarkable success for object detection and segmentation in the field of medical imaging [10,11,12]. In regional anesthesia and musculoskeletal US, attempts have been made to identify individual nerves, muscles and blood vessels using CNNs [13-18]. However, most of these studies are limited by small sized training datasets that affects their accuracy and generalizability.

In 2016, a dataset containing 5,365 manually annotated US images of the ISC was made publicly available by Kaggle to train machine learning models for identification of the brachial plexus [13]. This dataset was used by Baby et al.[14] and Kakade et al. [15] to train and validate a U-Net model [19] with image preprocessing steps and achieved dice coefficient scores of 0.71 and 0.69 respectively. In 2021, Wu et al.[16] adopted Region-aware pyramid aggregation model with Adaptive pyramid fusion to achieve a dice coefficient of 0.74 with real-time performance. In 2022, Ding et al.[17] provided a separate dataset of 1,052 manually annotated US images of ISC with best model performance achieving dice coefficient of 0.51 and a runtime of 5.2 FPS.

Ultrasonography is generally considered a challenging modality for image segmentation mainly because 1.) It generates poor contrast images which are dependent on operator skills for image acquisition as well as interpretation, 2.) The speckle noise characteristic to US imaging yields a granular texture in the images which reduces the overall SNR [20]. 3.) In addition, for automated real-time image segmentation, the hazy boundaries and speckle noise further introduce additional time-complexity hurdles on image preprocessing. Image segmentation on the brachial plexus sites also demands minimizing false positives, since inadvertent needle injury to the high stakes surrounding structures such as lung or blood vessels can be catastrophic.

Supraclavicular brachial plexus blockade, even though closely anatomically related to ISC brachial plexus, provides anesthetic coverage to a distinct area and is devoid of important side effects related to ISC block such as phrenic nerve paralysis and Horner syndrome [7]. However, despite its advantages and popularity among clinicians, SCBP segmentation has not been researched yet by the machine learning community. Furthermore, the two being closely related, there is a need to differentiate between these two adjacent structures for computer assisted diagnosis.

In the present study, we present an end-to-end DL approach for real time delineation of SCBP in neck US videos. This prospective study recruited 163 patients who were systematically scanned for SCBP using US. In addition, we also created a dataset of 33 ISC videos to evaluate the capability of the models to differentiate between SCBP and ISC. The resultant videos were then split into individual frames that were annotated by experienced anesthesiologists. A set of standard CNN models were trained as baseline models and their outputs were compared to the originally annotated images. In addition, two separate datasets of 15 videos each were constructed using different US scanners to test for generalizability of the model.

Contributions can be summarized as follows:

- 1- Build a large dataset of brachial plexus ultrasound videos with segmentation annotations done by experienced anesthesiologists for all the frames in each video. To the best of our knowledge, this is also the largest ultrasound video dataset in terms of number of frames annotated.
- 2 - Built models capable of differentiating the two adjacent nerve blocks - SCBP and ISC in real-time.
- 3 - Demonstrate generalizability of the model on data acquired through different ultrasound machines by different operators.

## **2.2. Methods**

### **2.2.1 Dataset curation**

This was a prospective study conducted jointly at the All India Institute of Medical Sciences, New Delhi, India and the International Institute of Information Technology, Hyderabad, India. This study follows the ‘Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research’ [21]. AIIMS Delhi Institutional ethics committee approval was obtained prior to commencement of the study (IEC/18/1/2020,RP-02/2020). Patients of age 20-80 years were included in this study. Exclusion criteria were patients known to have brachial plexus injury, brachial plexopathy, cervical trauma, limitations to neck movement or patient refusal. Patients were randomly approached in the preoperative holding area and were presented with written information regarding the study. After obtaining written informed consent, they were scanned with a linear high frequency probe of 12-15 MHz of US scanner eSaote MyLab One™(eSaote SPA, Genoa, Italy). Subjects were scanned on a 30-45 degree incline, with the head turned away from the side of interest with the overarching goal of focusing on the brachial plexus in the supraclavicular and interscalene space. US probes were then tilted, rotated and slid around in the cranio-caudal and medial to lateral directions to capture dynamic brachial plexus patterns in addition to adjoining areas of the neck such as muscles, fascia and blood vessels.

Random low, medium and high gain as well as different depth settings were used to increase the variation in the dataset. Table 1 summarizes the patient demographic characteristics. All personal identifiable information was cropped. The resulting videos were 542x562 pixel resolution.

Table 2.1. Data Set Patient Characteristics

<b>Characteristic</b>	<b>Total</b>	<b>Supraclavicular</b>	<b>Interscalene</b>
No. of Patients	196	157	33
No. of Videos*	196	157	33
Frames	34,926	29,129	4,622
Positive Data Frames	26,966	23,302	3,664
Gender			
Male	134 (68%)	107 (68%)	23(70%)
Female	62 (32%)	50 (32%)	10(30%)
Average Age	42 (13.6)	42 (13.5)	42 (13.9)
Average Height	170 (6.1)	171 (5.7)	169 (6.8)
Average BMI	25.3 (2.6)	25.4 (2.6)	25.2 (2.7)

Data is reported as n (%) or mean ( $\pm$ SD)

Positive data frames imply the presence of SCBP or ISC in a given frame.

6 videos contained only negative data and used only for training

## Annotation

A custom video annotation tool was built using OpenCV software [22] for the purpose of annotation, details of which have been uploaded on the GitHub website [23]. In a given video, a random frame was annotated for SCBP by an expert anesthesiologist using bounding boxes and was designated as seed frame. The bounding boxes were drawn to cover the approximate SCBP area and were not meant to annotate individual roots. In addition, boxes were drawn keeping in mind to stay clear of critical structures such as the subclavian artery and pleura in all frames. Subsequent frames were automatically annotated using the CSR-DCF and KCF object-tracking algorithms [24,25]. Each automated annotation was either approved or rejected by the annotator. If rejected, the bounding boxes were redrawn as new seeds for automated annotations. The individual boxes were then fused to create a single mask. The contour of this mask was then shrunk to better match the boundaries of the plexus using morphological geodesic active contour algorithm [26]. The parameters of the algorithm were chosen for each individual video separately and are available in the dataset. The final shrunken masks were labeled as “ground truth” masks (Fig. 1) which were used to train the neural network. The frames in which the primary annotator was uncertain, a second expert was called in to review the frames. If still uncertain of the presence or absence of SCBP or ISC, these frames were discarded to avoid noisy annotations.

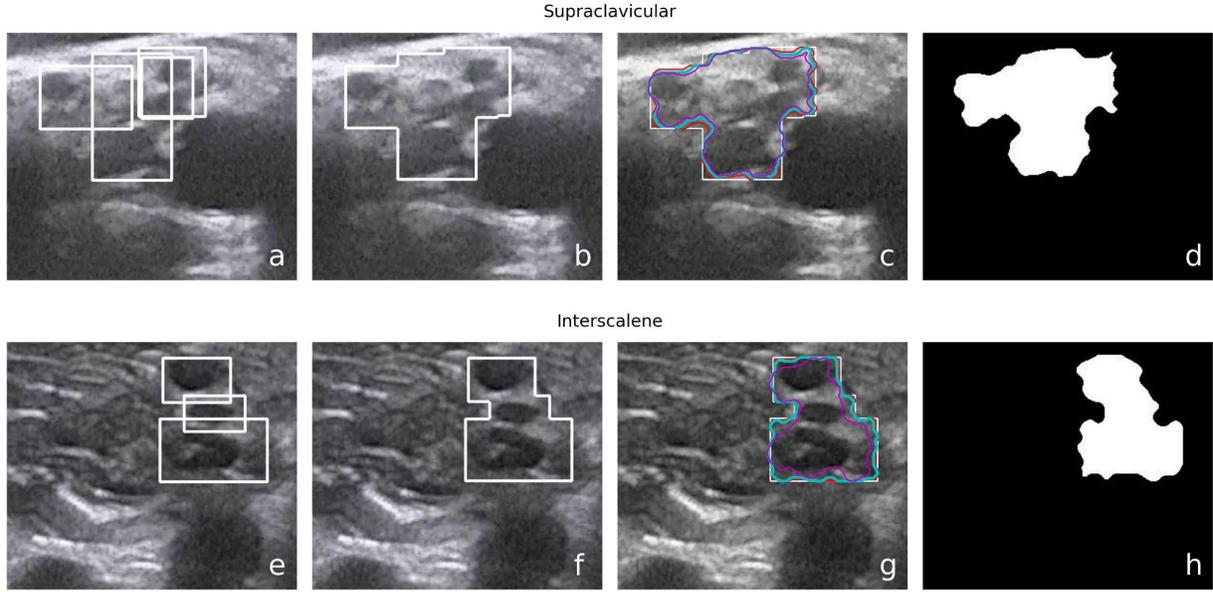


Fig 2.1. a,e) Hand drawn annotation bounding boxes to cover the plexus as a whole while avoiding critical structures such as the pleura and subclavian artery b,f) outer boundary of the union of boxes selected as the initial contour c,g) Various shrunken contours proposed by the morphological geodesic active contour algorithm using different parameters d,h) Best contour approximating the boundaries of the plexus chosen by experienced anaesthesiologists as the final ground truth

---

**Algorithm 2.1** CSR-DCF object-tracking algorithm [24]

---

**Require:**

Image  $I_t$ , object position on previous frame  $\mathbf{p}_{t-1}$ , scale  $s_{t-1}$ , filter  $\mathbf{h}_{t-1}$ , color histograms  $\mathbf{c}_{t-1}$ , channel reliability  $\mathbf{w}_{t-1}$

**Ensure:**

Position  $\mathbf{p}_t$ , scale  $s_t$  and updated models.

**Localization and scale estimation:**

- 1: New target location  $\mathbf{p}_t$ : position of the maximum in correlation between  $\mathbf{h}_{t-1}$  and image patch features  $\mathbf{f}$  extracted on position  $\mathbf{p}_{t-1}$  and weighted by the channel reliability scores.
- 2: Using location  $\mathbf{p}_t$ , estimate new scale  $s_t$ .

**Update:**

- 3: Extract foreground and background histograms  $\mathbf{c}^f, \mathbf{c}^b$
  - 4: Update foreground and background histograms  

$$\mathbf{c}_t^f = (1 - \eta_c)\mathbf{c}_{t-1}^f + \eta_c\mathbf{c}^f, \mathbf{c}_t^b = (1 - \eta_c)\mathbf{c}_{t-1}^b + \eta_c\mathbf{c}^b$$
  - 5: Estimate reliability map  $\mathbf{m}$
  - 6: Estimate a new filter  $\tilde{\mathbf{h}}$  using  $\mathbf{m}$  (constrained filter optimization)
  - 7: Estimate channel reliability  $\tilde{\mathbf{w}}$  from  $\mathbf{h}$
  - 8: Update filter  $\mathbf{h}_t = (1 - \eta)\mathbf{h}_{t-1} + \eta\tilde{\mathbf{h}}$
  - 9: Update channel reliability  $\mathbf{w}_t = (1 - \eta)\mathbf{w}_{t-1} + \eta\tilde{\mathbf{w}}$
-

### 2.2.2 Data augmentations

Four data augmentations with random parameters were employed to enrich variations in the training data.

1. Horizontal flip: Images were horizontally flipped with a 50% probability to mimic the left-right dexterity in SCBP sono-anatomical patterns.
2. Rotation: Images were rotated by a random angle between -10 and 10 degrees to mimic the incline-angle variations of the US probe.
3. Gamma intensity transformation: The image intensities were transformed with random gamma values to mimic variations in US frequency gain (Fig. 2.2). The original high gain images were transformed with a random gamma value between 1.5 and 2.0 to mimic medium and low gain images. The original low gain images were transformed with a random gamma value between 0.5 and 0.75 to mimic high and medium gain images. The original medium gain images were transformed with a random gamma value between 0.75 and 1.33 to mimic high and low gain images.

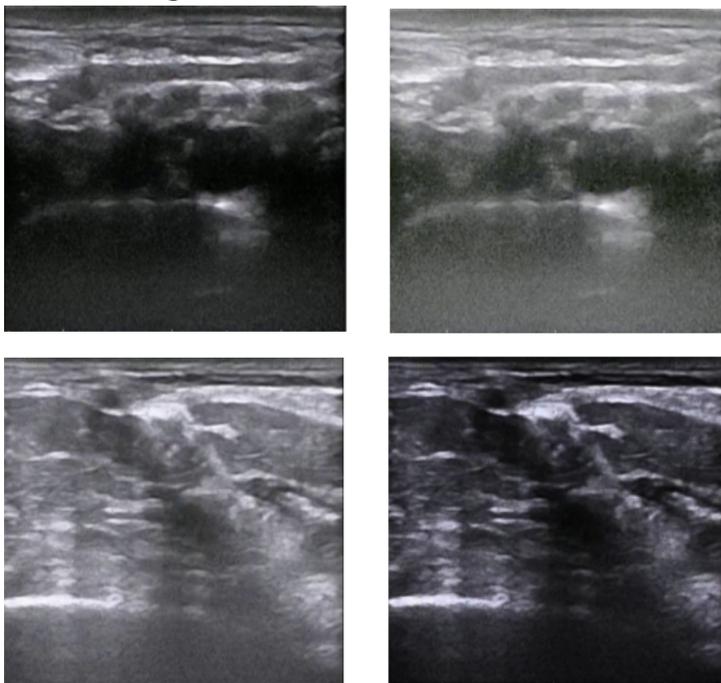


Fig. 2.2 Gamma Intensity Transformation:

Left Column: Original images; Right Column: Gamma transformed images;

Top Row: Low gain image transformed with gamma value of 0.5;

Bottom Row: High gain image transformed with gamma value of 2.0

4. Gaussian synthetic acoustic shadows: Ultrasound can create acoustic shadows because air or tissue inhibiting sound waves from penetrating deeper into the body. This effect is mimicked using a 2D Gaussian multiplication to the image.

A synthetic shadow  $G(x,y)$  centered at a random pixel  $(p_x, p_y)$  is generated using a 2D Gaussian as shown in below equation, where the standard deviations,  $\sigma_x$  and  $\sigma_y$  determine radius of the shadow and  $i$  determines the intensity of the shadow

$$G(x, y) = 1 - ie \left[ -\frac{(x-p_x)^2}{2\sigma_x^2} - \frac{(y-p_y)^2}{2\sigma_y^2} \right]$$

Augmented Image  $I_a$  is obtained by multiplication of  $G(x,y)$  above to the image  $I(x,y)$  (Fig. 2.3)

$$I_a(x, y) = I(x, y) * G(x, y)$$

The original images were transformed using Gaussians shadows having random standard deviations between  $\frac{1}{8}$  and  $\frac{1}{4}$  times image size and random shadow strength intensities between 0.4 and 0.8.

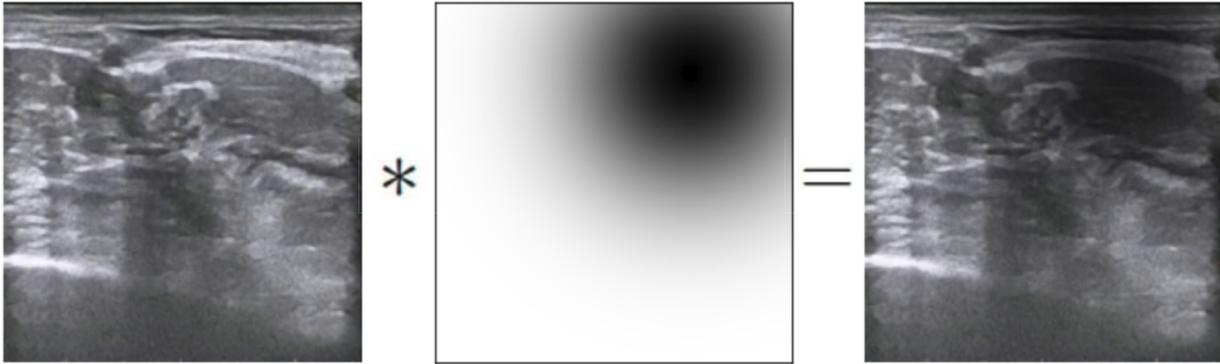


Fig. 2.3 Left: Original image; Center: Synthetic acoustic shadow using a 2-D Gaussian of standard deviation of  $\frac{1}{4}$  times image size and 0.5 shadow strength intensity; Right: Transformed image with an acoustic shadow

### 2.2.3 Neural Network

From a wide range of available off-the-shelf segmentation models, we shortlisted two models - DeepLabv3 [28] and UNet [19], and their enhancements - DeepLabv3+ [29] and UNet++ [30] based on established usage in medical image segmentation literature [31, 32], real-time processing capability and low memory usage. All the four models use 50-layer residual-net as backbone [33].

One of the limitations while using traditional fully convolutional neural networks for image segmentation tasks is that the spatial resolution is progressively downsampled resulting in information loss and low resolution output masks with fuzzy object boundaries. The DeepLabV3 model addresses this by using Atrous convolutions [34] and Atrous Spatial Pyramid Pooling (ASPP) [35-37] which help in extracting information with a wider receptive field and thus avoiding the loss of spatial resolution. The UNet model

mitigates the downsampled resolution problem through up-convolutions on the expanding path and concatenations with features from the downsampling path, resulting in a symmetrical U shaped network.

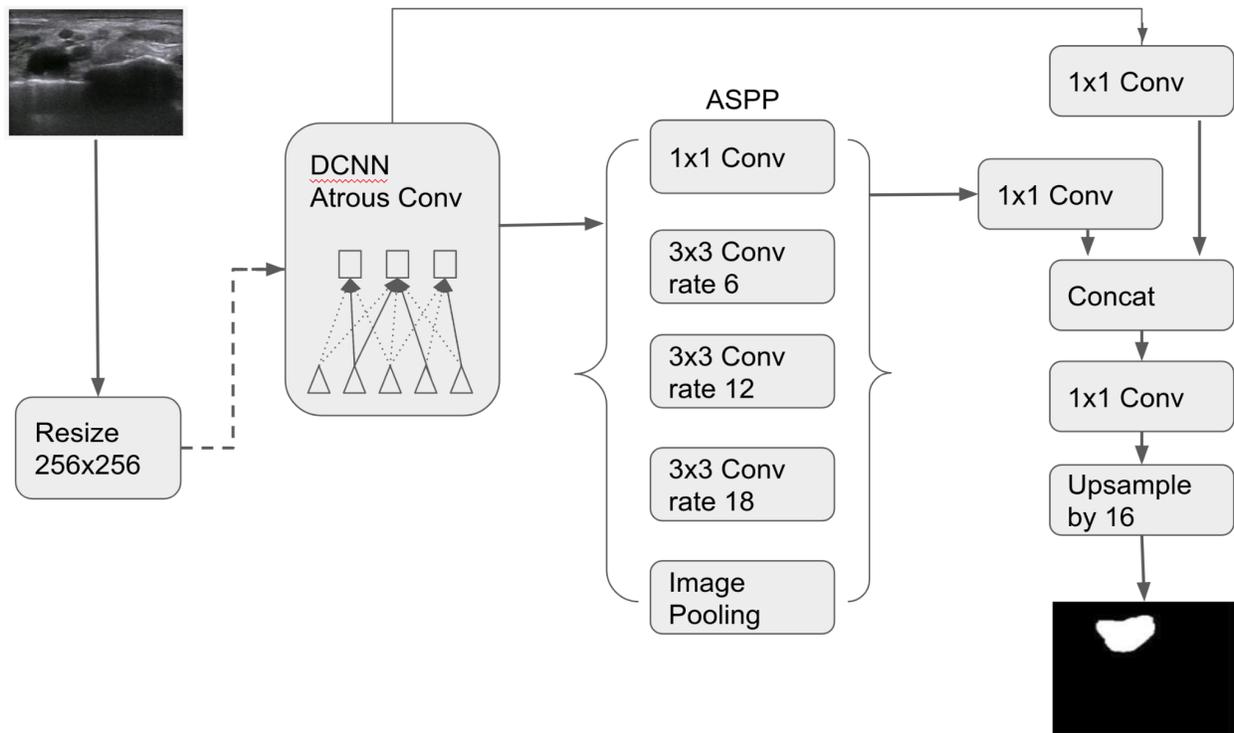


Fig 2.4. Framework of DeepLabV3 architecture

**Atrous convolution:** For a 2D feature map, for each location  $i$  on the output  $y$  and a kernel  $w$ , atrous convolution is implemented over the input  $x$ :

$$y[i] = \sum_k x[i + r \cdot k]w[k]$$

where the atrous rate  $r$  corresponds to the stride sampling the input signal, equivalent to convolving the input  $x$  with upsampled kernels produced by inserting  $r - 1$  nulls between two successive kernel values along each spatial dimension (named atrous convolution after the French word trous). Standard convolution is a special case for rate  $r = 1$ , and atrous convolution allows for dynamic modification of the kernel's field-of-view by changing the rate value.

Model design and training was done in the Python programming language using the PyTorch 1.8 deep learning framework [38]. The models were initialized with pretrained Imagenet weights [39] and were trained using Adam optimizer [40] for 25-50 epochs with early stopping and batch size 16 on an Ubuntu 20.04 machine with an NVIDIA RTX 3090 GPU. Training took approximately 9 minutes per epoch and minimized pixel-level binary cross-entropy loss.

### 2.2.4 Internal validation

For internal model validation, stratified 5-fold cross-validation was performed. The training:validation:test proportions were 61:19:20 for each fold. The stratified samples were chosen to keep roughly equal proportions for dexterity (left/right), frequency gains (high/medium/low), and gender (male/female) in each of the training, validation and test sets.

### 2.2.5 Generalizability

For testing generalizability of the model, two separate datasets of 15 videos each were constructed using different US machines. The first dataset was created using Sonosite M-Turbo™ (Fujifilm Sonosite Inc, Bothell, WA, USA) which is a cart based US scanner that produces videos of similar quality to the original machine in the main dataset. Six of the 15 videos were acquired with needle in-situ to test the performance of the model in the presence of needle. These 15 videos contained a total of 2,186 frames.

The second dataset was created using a relatively inexpensive portable handheld machine Butterfly-IQ™ (Butterfly Network Inc., Burlington, MA, USA) which produces videos of lower quality [41,42]. The 15 videos from the second machine contained a total of 3,088 frames.

Models were evaluated on these datasets with and without fine-tuning.

### 2.2.6 Evaluation

The goal of this study is not pixel-accurate segmentation but rather to identify the approximate location of the plexus in the US images. Therefore, the performance evaluation was done using both object detection metrics and segmentation metrics.

For object detection metrics, the precision, recall and F-1 scores were calculated using intersection over union (IoU) thresholds [43]. The IOU measures the area overlapping between the prediction mask and the ground truth divided by the area of union between them. The IOU is then compared to a given threshold ( $t$ ) depending on which, the predictions are classified as correct or incorrect. If the  $\text{IoU} \geq t$ , then the prediction was considered correct or true positive. If  $\text{IoU} < t$ , the detection was considered as incorrect or false positive. Failure to detect ground truth in a given frame was considered false negative. Smistad et. al used a 25% IOU threshold to delineate nerves and blood vessels in the axillary space stating a goal of only highlighting the approximate location of nerves and not pixel perfect nerve boundary delineation [18]. In computer vision literature, 50% IOU is used as a standard threshold for classifying the presence or absence of an object in an image. The IOU thresholds tested in the current study were set at 25% and 50%. The predictions too small to be regarded as SCBP or ISC were discarded. The threshold for this was set at 20% of the median area occupied by the ground truth in the entire dataset, i.e; 3240 pixels for median SCBP area and 914 pixels for ISC in resized 256x256 images.

The different spatial resolutions of different US machines were brought to common ground of width 256 pixels for comparison. The original dataset and Butterfly-IQ™ dataset were evaluated at 256x256 resolution and Sonosite M-Turbo™ dataset were evaluated at 256x192 resolution.

For segmentation metrics, dice coefficient was calculated in three different ways:

1. Dice coefficient for both plexuses in all frames of all videos: This measured the ability of the model to differentiate the two plexuses. When the model correctly identified the absence of nerve in any image i.e. in the true-negative case, 100% dice-score was assigned for that image. However, this inflated the dice coefficient.
2. Dice coefficient for SCBP was calculated using only SCBP videos and dice coefficient for ISC was calculated using only ISC videos. Here also, the frames with true-negative cases inflated the overall dice scores for the video.
3. Dice coefficients were calculated separately for SCBP and ISC using only the positive frames in SCBP and ISC videos respectively. The frames with true-negative cases were ignored and not taken for consideration.

For all the above metrics, the average metric of all considered videos is the final metric. The metric for individual video is the average metric of all frames in that video.

### **2.3. Results**

For a total of 163 videos, six did not contain SCBP and were used for training purposes only. In total, 29,129 image frames were annotated with 23,302 positive frames (SCBP present) and 5,827 negative frames (SCBP absent). Prediction runtime of the algorithm was within real-time constraints. The average run-rate was measured to be more than 30 fps using the NVIDIA RTX 3090 GPU.

Table 2, shows the average precision, recall and F score  $\pm$  standard deviation at 25% and 50% IOU thresholds, and three dice coefficients as described in the Evaluation section for the four baseline models' performance on SCBP and ISC.

Table 2.2: Comparison of baseline models on Object Detection & Segmentation metrics

	DeepLabV3	DeepLabV3+	UNet	UNet++
<b>Supraclavicular BP (SCBP)</b>				
Object Detection at 25% IOU				
Precision	0.93 (0.17)	0.94 (0.15)	0.94 (0.16)	0.95 (0.16)
Recall	0.96 (0.07)	0.97 (0.13)	0.96 (0.07)	0.97 (0.08)
F-1 score	0.94 (0.16)	0.95 (0.13)	0.95 (0.15)	0.96 (0.15)
Object Detection at 50% IOU				
Precision	0.86 (0.22)	0.87 (0.23)	0.87 (0.21)	0.88 (0.21)
Recall	0.95 (0.11)	0.96 (0.06)	0.94 (0.09)	0.96 (0.10)
F-1 score	0.90 (0.20)	0.91 (0.20)	0.90 (0.19)	0.92 (0.19)
Dice Coefficient				
All videos	0.81 (0.12)	0.83 (0.12)	0.83 (0.11)	0.84 (0.11)
SCBP videos	0.78 (0.10)	0.80 (0.10)	0.79 (0.09)	0.81 (0.09)
SCBP positive frames	0.74 (0.10)	0.76 (0.10)	0.75 (0.09)	0.77 (0.09)
<b>Interscalene BP (ISC)</b>				
Object Detection at 25% IOU				
Precision	0.76 (0.44)	0.78 (0.32)	0.79 (0.33)	0.80 (0.36)
Recall	0.88 (0.16)	0.90 (0.22)	0.84 (0.37)	0.88 (0.26)
F-1 score	0.82 (0.44)	0.84 (0.31)	0.81 (0.34)	0.84 (0.36)
Object Detection at 50% IOU				
Precision	0.57 (0.36)	0.59 (0.31)	0.59 (0.35)	0.60 (0.36)
Recall	0.85 (0.28)	0.87 (0.23)	0.79 (0.36)	0.81 (0.34)
F-1 score	0.68 (0.36)	0.70 (0.32)	0.68 (0.37)	0.69 (0.36)
Dice Coefficient				
All videos	0.93 (0.15)	0.94 (0.16)	0.93 (0.15)	0.94 (0.15)
ISC videos	0.64 (0.19)	0.69 (0.19)	0.65 (0.19)	0.68 (0.19)
ISC positive frames	0.59 (0.20)	0.64 (0.19)	0.61 (0.20)	0.63 (0.19)

Data Reported as mean( $\pm$ SD)  
 All models had ResNet-50 backbone

Figure 2.5 depicts the precision-recall curves at 25% and 50% IOU thresholds for UNet++ model on SCBP and ISC.

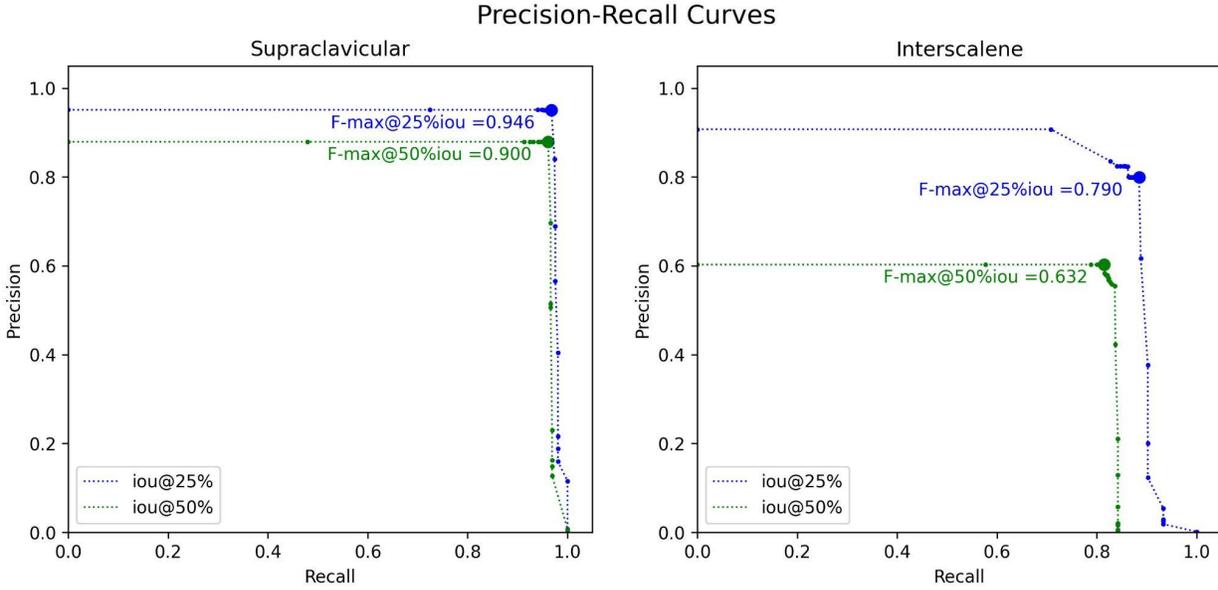


Fig. 2.5 Precision-Recall Curves for IoU overlap criteria of 25% and 50%

Table 2.3 shows the performance of UNet++ model before and after fine-tuning on the datasets created through Sonosite M-Turbo™ which is a similar quality machine and a lower quality handheld machine i.e. Butterfly IQ™ scanner.

Table 2.3: UNet++ performance on different machines

	Sonosite M-Turbo™		Butterfly IQ™	
	Before FT	After FT	Before FT	After FT
<b>Object Detection at 25% IOU</b>				
Precision	0.99 (0.03)	0.99 (0.02)	0.77 (0.19)	0.96 (0.02)
Recall	0.96 (0.06)	0.99 (0.04)	0.76 (0.20)	0.98 (0.03)
F-1 score	0.97 (0.04)	0.99 (0.03)	0.76 (0.19)	0.97 (0.03)
<b>Object Detection at 50% IOU</b>				
Precision	0.80 (0.24)	0.95 (0.05)	0.23 (0.26)	0.60 (0.25)
Recall	0.90 (0.09)	0.98 (0.04)	0.42 (0.40)	0.96 (0.06)
F-1 score	0.85 (0.19)	0.96 (0.04)	0.30 (0.30)	0.74 (0.19)
<b>Dice Coefficient</b>				
All frames	0.69 (0.09)	0.78 (0.03)	0.55 (0.17)	0.67 (0.08)
Positive frames	0.69 (0.08)	0.77 (0.03)	0.45 (0.16)	0.66 (0.08)

FT: Fine-Tuning  
Data is reported as n(%) or mean(±SD)

Supplementary video 2.1 shows model performance with comparison to ground truth.

Supplementary video 2.2 illustrates model predictions differentiating SCBP and ISC with US probe movement along the neck.

Supplementary video 2.3 shows model interaction with needle insertion.

## **2.4. Discussion**

In the present work, we demonstrate a DL based approach that successfully delineates and tracks the SCBP across the given US videos of the neck with high accuracy and reproducibility. The prediction rate was greater than 30 fps, which is well above the real-time constraints. When applied to an unseen separate dataset constructed using a separate US scanner of similar quality, the model performance remained consistent. To the best of our knowledge, this is the first DL based approach trained on a large dataset that delineates SCBP in US videos with high precision and reliability. In addition, the model trained jointly on SCBP and ISC was able to successfully differentiate between SCBP and adjacent ISC as shown in Supplementary video 2.2.

Over the past two decades, ultrasonography has rapidly ascended to become indispensable in clinical practice due to its unique set of advantages, such as absence of ionizing radiation, portability and ability to produce real-time imaging which can be used for both diagnostics and interventions. With continuous advancement in computational power to give out faster results, the idea of leveraging DL to augment US experience is getting closer to realization, wherein algorithm segmentations can be superimposed upon US videos in real time that may assist clinicians in making decisions. However, for a robust training of DL algorithms and to achieve desired performance, a large training dataset is usually needed, that needs to be manually annotated by experts, often proving to be a laborious task. A common roundabout is thus taken where human experts annotate one or two US images per clip to create a small to moderate sized ground truth dataset. This approach leaves DL models at a disadvantage because it does not incorporate the dynamic patterns of neural structures, as in our case SCBP and ISC, that continually change in shape and size with the US probe movement. Furthermore, transfer learning, a technique used to compensate for small amounts of training data is commonly applied wherein pre-trained models that identify real world natural image objects are further trained on small sized medical image datasets [44, 45]. Although successful to some extent, this leads to poor generalization to US images as the speckle noise seen in grayscale US images is not present in natural images. Also, the structures in US images tend to have fuzzy boundaries compared to well-defined boundaries seen in natural image objects. In the current study,

we partially automated annotation using object tracking and active contour algorithms that helped anesthesiologists annotate a seemingly large image dataset in a comprehensive manner within a short period of time. Scanning protocols were prospectively laid out to capture variable ISC and SCBP patterns seen with different US probe orientations. Instead of using one or two images per video, sequential frames were annotated in different depth and gain settings for exhaustive training of our CNNs that we believe greatly adds to the accuracy and robustness of our approach.

Upon review of literature, various authors have attempted to delineate brachial plexus in ultrasound images at the interscalene and axillary spaces using machine learning. Utilizing the 5,365 training images of Kaggle dataset, Baby et al. [14] and Kakade et al. [15] implemented the UNet model to achieve dice coefficient of 0.71 and 0.69 respectively, but evaluation on the test dataset could not be done since the test dataset annotations were not released by Kaggle. Both of these approaches utilized time consuming image-processing steps that made the real-time inference infeasible. Wu et al improved upon this with Region-aware pyramid aggregation model and Adaptive pyramid fusion to achieve a dice coefficient of 0.74 at real-time performance [16]. Ding et al. [17] created a new data set of 1,052 frames from 101 videos and publicly released their training dataset of 955 frames from 91 videos containing annotations for ISC as well as surrounding anatomical structures i.e. muscles and blood vessels. They improved upon Mask-RCNN through a spatial local contrast feature extraction module and self-attention gate thereby utilizing the spatial information of surrounding anatomical structures and achieved a dice coefficient of 0.51 at 5.19 FPS. In comparison, a standard of-the-shelf UNet++ model trained on our 3,664 ISC frames from 33 videos achieved a dice coefficient of 0.68 at more than 30 FPS. However, the higher score in our study could be attributed to different image characteristics and bigger dataset in our study. Smistad et al.[18] used U-Net to delineate individual brachial plexus nerves in axillary space and reported an F-score ranging from 0.39 (radial nerve) to 0.73 (median nerve) using 25% IoU overlap criterion with real-time inference capability. Our results for SCBP showed a superior F-score of 0.95 and 0.90 for 25% and 50% IOU threshold respectively.

It is noteworthy that none of the abovementioned studies evaluated their models for generalizability by testing on a separate dataset curated from other US scanners. This could be an important limitation since different US scanners tend to have different image characteristics. The present study tested its model on two separate scanners - a traditional cart based Sonosite M-Turbo™ with superior image quality (Fujifilm Sonosite Inc, Bothell, WA, USA) and a hand held portable Butterfly IQ scanner that can be plugged into smartphones and run with an app with a compromise of a lower image quality. Our results showed a dice coefficient of 0.69 and 0.55 for delineating SCBP on images generated from Sonosite and Butterfly US

scanners respectively, without any additional fine-tuning . The dice coefficient scores further improved to 0.78 and 0.67 respectively after fine-tuning.

The present study has its limitations. Our CNN was trained from a set of US videos obtained by four anesthesiologists trained in regional anesthesia according to a well defined protocol focusing only on two specific areas of the patient's neck, i.e; the supraclavicular and interscalene spaces using one US scanner. Variations may come to model performance in diverse clinical settings with different operators using different US scanners as depicted in Table 3. That being said, performance on a new machine or patient population can be improved with fine-tuning on a small amount of annotated data from the concerned machine or population . Rapid and jerky probe movements may result in erroneous predictions that tend to improve when the probe movements are slowed and stabilized over one area. This is because of the motion artifacts that cause blurring of structures leading to erroneous interpretations. The demographic characteristics of our training set consisted mostly of average built subjects which may again lead to variable model performance when used on extremes of body habitus.

As an extension to our scanning protocol, a few videos were captured to assess model performance with needle in situ while performing nerve block as shown in supplementary video 3. The model performance remained consistent in the presence of the needle but dropped significantly when the needle entered the confines of the plexus. This may be explained as the entry of the needle changes the plexus anatomy that has not been incorporated in the training of models. This finding underscores the need for future studies and datasets that include videos of PNB being performed with needle in situ to train models that can further improve their performance in such conditions.

One of the major obstacles for DL based studies is acquiring a high fidelity dataset containing a well labeled ground truth that experts agree upon. As a resource for future studies to train and validate their models, we are releasing our full dataset of 227 videos consisting of 41,000 annotated images of SCBP and ISC that can be accessed on the GitHub website [23]. To the best of our knowledge, this is one of the largest datasets of US images annotated by human experts for SCBP and ISC that could help future DL studies on ultrasonography. Our combined SCBP and ISC US video dataset further complements previous available datasets which have consisted only of isolated US video frames. We hope that further video based models that utilize temporal information can be developed by the research community using this dataset.

In conclusion, in this chapter we presented a DL based approach for real time delineation of SCBP in US videos with high accuracy and reliability. This model may be incorporated in the US scanners that could potentially assist regional anesthesia trainees to administer US-guided SCBP nerve block. We also believe the contribution of an open labeled dataset will greatly enhance the development of new deep learning applications in regional anesthesiology and musculoskeletal ultrasonography.

Dataset and code availability:

The dataset and code used for annotation in this work can be freely downloaded from:

<https://github.com/Regional-US/>

## *Chapter 3*

# **Needle Guidance for Autonomous Robotic Ultrasound Guided Regional Anesthesia**

### **3.1 Introduction**

Ultrasound guided regional anesthesia (UGRA) involves approaching target nerves through a needle in real-time, enabling precise deposition of drug with increased success rates and fewer complications [3]. Development of autonomous robotic systems capable of administering UGRA is desirable for remote settings and localities where anesthesiologists are unavailable [47]. Real-time segmentation of nerves, needle tip localization and needle trajectory extrapolation are required for developing such a system.

#### **3.1.1 Ultrasound Challenges**

Ultrasonography is generally considered a challenging modality for image segmentation because of poor contrast images, operator dependency, and the reduced SNR by granular image texture from speckle noise. In addition, for real-time segmentation, the hazy boundaries and speckle noise introduce additional time-complexity hurdles on image preprocessing. Needles occupy very few pixels in images which pose additional challenges as their appearance is often degraded by reverberation, comet tail, side-lobe, beam-width, or bayonet artifacts. If the needle does not lie parallel to the ultrasound beam's plane, it is visible only partially at the intersection with the plane. Also, the segmentation process may be disturbed

by adjacent linear anatomical structures (e.g., borders of bones or fasciae pleura etc). The invasive nature of needles limits the possibility of acquiring large scale data required for deep learning methods.

- **Common artifacts** are always present in ultrasound images because they are caused by the physical limitations of ultrasound imaging.
- **Uncommon artifacts** only occasionally distort ultrasound images, but can be very confusing to an inexperienced operator if they do occur. These artifacts are often caused by technical problems or by the patient's anatomy.

The common artifacts are generally easy to identify and do not usually interfere with the diagnosis. However, the uncommon artifacts can be very confusing, especially for an inexperienced operator [46].

### **Needle related artifacts**

- The difference between real and ultrasonic appearance of a needle can be due to one or more of the three physical reasons: spatial resolution, beam reflection, or acoustic shadowing. Stainless steel needles and bones have high acoustic impedance and appear as bright line on ultrasound images. This is why metal needles can sometimes be mistaken for bone on ultrasound images.
- **Reverberation artifact** occurs when sound waves bounce back and forth between two reflective surfaces of needle's posterior and anterior cylindrical walls. This can create a series of echoes on the ultrasound image, which can obscure the underlying structures.
  - **Comet tail artifact**: This is a long, thin echo that trails behind a strong reflector caused by the reflection of sound waves from small, highly reflective objects. The tail might appear penetrating deep into the tissue but the actual needle is much above at the upper limit of the tail.
- Beam-width artifact and side-lobe artifact are two common ultrasound artifacts that can occur when the ultrasound beam is not perfectly focused.
  - **Beam-width artifact** occurs when the ultrasound beam is wider than the structure being imaged causing the edges of the structure to appear blurred or fuzzy.
  - **Side-lobe artifact** occurs when sound waves are emitted from the sides of the ultrasound transducer. These sound waves can reflect off of surrounding tissues and create faint echoes on the ultrasound image.
- The sloped part of obliquely cut needle tips can generate a small line slightly ahead and below the needle's predominant line.

- **Bayonet artifact:** When an ultrasound needle passes through or under tissues with different acoustic impedances, the reflected beams from different parts of the needle but at the same depth arrive back to the probe at slightly different times. The ultrasound machine then calculates the depth of the needle based on the time taken by the sound waves to return, but this calculation is incorrect because the sound waves from different parts of the needle arrived at different times. This can cause the needle to appear bent on the ultrasound image, even though it is actually straight.

### 3.1.2 Recent work

In recent years, attempts have been made to identify individual nerves, surrounding structures and needles using object detection and segmentation neural networks [48]. However, most of these studies are limited by time complexity and small-sized training datasets that affects their accuracy and generalizability. Most of the automatic needle detection methods are based on phantoms or ex-vivo porcine and chicken tissues [48]. However, compared to phantoms, the human tissues are much more heterogeneous with many artifacts, underscoring the need to work with real in-vivo human data.

### 3.1.3 Contribution

In this chapter, we 1) define automated anesthesia needle targets derived from CNN nerve segmentation predictions 2) design and implement automated needle segmentation, needle tip localization and needle trajectory extrapolation. We also publicly release needle annotations of 6 ultrasound SCBP videos comprising of 1,061 frames.

To the best of our knowledge, this is the largest publicly released dataset of annotated needle frames in ultrasound videos created from human subjects. This is also the first study on simultaneous identification of the needle and its regional anesthesia target in given nerve ultrasound videos.

## 3.2 Materials and Methods

### 3.2.1 Dataset and Annotation

Out of the 227 videos from Chapter 3, six videos consisting of 1,061 frames containing both the needles and SCBP were used. SCBP annotation method is presented in Chapter 3. The coordinates of the needle tip and needle tail-end at the skin were annotated by experienced anesthesiologists. Four pixel thick white lines between these coordinates were drawn to create needle masks. Ultrasound videos, nerve masks and needle masks are publicly available at github website.

### 3.2.2 Nerve block Target

The SCBP lies adjacent to the subclavian artery which can be identified due to its pulsating nature. Underneath the plexus, lies the clavicle bone and pleura or the outer layer covering the lungs. A successful block involves charting the needle course to deposit the drug below and above the plexus as well as preventing any injury to the adjacent artery and the pleura.

In order to define the needle target, first an ellipse is fitted on the contour of the nerve segmentation prediction (Fig 3.1). The top and bottom peripheries of the nerve plexus are defined as elliptic arcs between  $\pm 30$  degrees from a vertical line through the center of the ellipse regardless of the orientation of the ellipse. Finally, the anesthesia target is defined as the region 4 pixels (0.8 mm) above and below these elliptic arcs as shown in Fig. 3.2

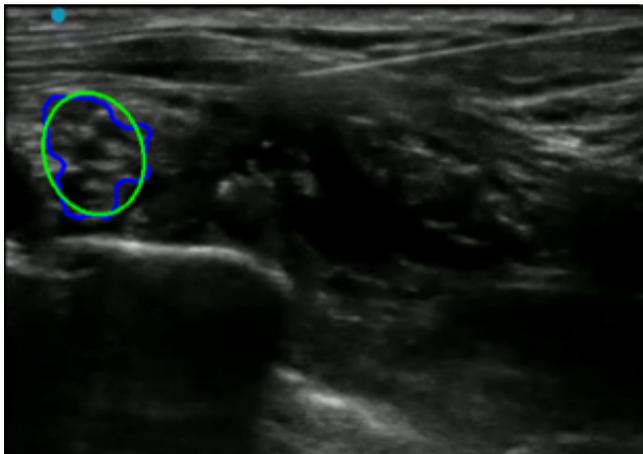


Fig. 3.1 An ellipse (green) is fitted on the contour (blue) of the nerve segmentation prediction

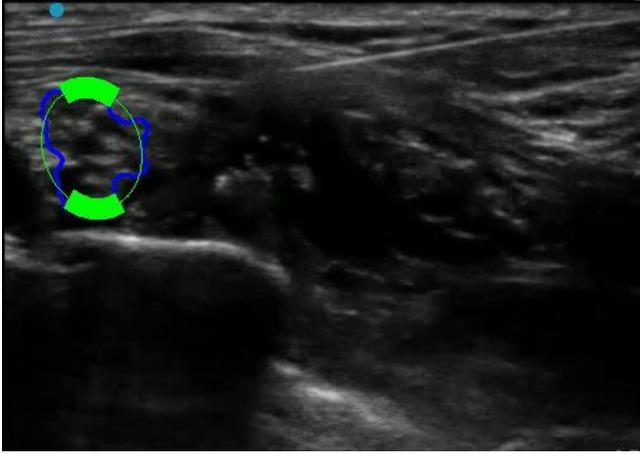


Fig.3.2. Needle target (green arcs)

### 3.2.3 Neural Network

A pre-trained (on natural RGB images) UNet++ network [30] was first fine-tuned on a large dataset of 196 ultrasound videos wherein the network learned the ultrasound domain and to segment SCBP and ISC as two output channels. In the next step, a second fine-tuning on five needle videos was done to adapt the second output channel of ISC to predict needle mask. After training on five random videos we tested the needle in the sixth video.

### 3.2.4 Needle Trajectory

The problem of partial visibility of the needle was overcome by trajectory extrapolation using the visible part. Radon transform was applied on the network's needle mask prediction to obtain a sinogram. This transform essentially comprises projections at different angles.

#### Radon Transformation [27]

Radon transform is the integral transform which takes a function  $f$  defined on the plane to a function  $Rf$  defined on the (two-dimensional) space of lines in the plane, whose value at a particular line is equal to the line integral of the function over that line. If a function  $f$  represents an unknown density, then the Radon transform represents the projection data obtained as the output of a tomographic scan. Hence the inverse of the Radon transform can be used to reconstruct the original density from the projection data, and thus it forms the mathematical underpinning for tomographic reconstruction.

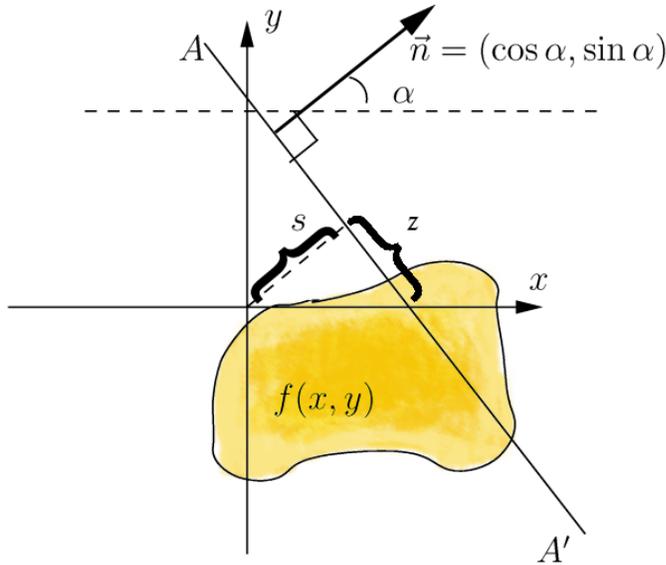


Figure 3.3 Radon transform. Credit: Begemotv2718. Wikimedia Commons under CC BY-SA 3.0

**Definition:**

Let  $f(x,y)$  be a continuous function that satisfies:

- the double integral  $\iint \frac{|f(x,y)|}{\sqrt{x^2+y^2}} dx dy$  extending over the whole plane, converges
- for any arbitrary point  $(x,y)$  on the plane it holds that

$$\lim_{r \rightarrow \infty} \int_0^{2\pi} f(x + r \cdot \cos\varphi, y + r \cdot \sin\varphi) d\varphi = 0$$

The Radon transform,  $Rf$ , is a function defined on the space of straight lines  $L \subset \mathbb{R}^2$  by the line integral along each such line as:

$$Rf(L) = \int_L f(x)|dx|$$

Concretely, the parametrization of any straight line  $L$  with respect to arc length  $z$  can always be written:

$$(x(z), y(z)) = (z \cdot \sin\alpha + s \cdot \cos\alpha, -z \cdot \cos\alpha + s \cdot \sin\alpha)$$

where  $s$  is the distance of  $L$  from the origin and  $\alpha$  is the angle the normal vector to  $L$  makes with the X-axis. It follows that the quantities  $(\alpha, s)$  can be considered as coordinates on the space of all lines in  $\mathbb{R}^2$ , and the Radon transform can be expressed in these coordinates by:

$$\begin{aligned} Rf(\alpha, s) &= \int_{-\infty}^{\infty} f(x(z), y(z)) dz \\ &= \int_{-\infty}^{\infty} f((z \cdot \sin\alpha + s \cdot \cos\alpha), (-z \cdot \cos\alpha + s \cdot \sin\alpha)) dz \end{aligned}$$

More generally, in the  $n$ -dimensional Euclidean space  $\mathbb{R}^n$ , the Radon transform of a function  $f$  satisfying the regularity conditions is a function  $Rf$  on the space  $\sum_n$  of all hyperplanes in  $\mathbb{R}^n$ . It is defined by:

$$Rf(\xi) = \int_{\xi} f(x) d\sigma(x), \quad \forall \xi \in \sum_n$$

where the integral is taken with respect to the natural hypersurface measure,  $d\sigma$  (generalizing the  $|dx|$  term from the 2-D case). Observe that any element of  $\sum_n$  is characterized as the solution locus of an equation  $x \cdot \alpha = s$ , where  $\alpha \in S^{n-1}$  is a unit vector and  $s \in \mathbb{R}$ . Thus the  $n$ -dimensional Radon transform may be rewritten as a function on  $S^{n-1} \times \mathbb{R}$  via:

$$Rf(\alpha, s) = \int_{x \cdot \alpha = s} f(x) d\sigma(x)$$

It is also possible to generalize the Radon transform still further by integrating instead over  $k$ -dimensional affine subspaces of  $\mathbb{R}^n$ . The X-ray transform is the most widely used special case of this construction, and is obtained by integrating over straight lines.

### Trajectory

The Radon transform data is often called a sinogram because the Radon transform of an off-center point source is a sinusoid. Assuming that the sinogram's maximum value corresponds to the needle trajectory's angle and that the lowest point in the needle mask is the needle tip, a line is drawn from the needle tip to show the needle trajectory (Fig. 3.4).

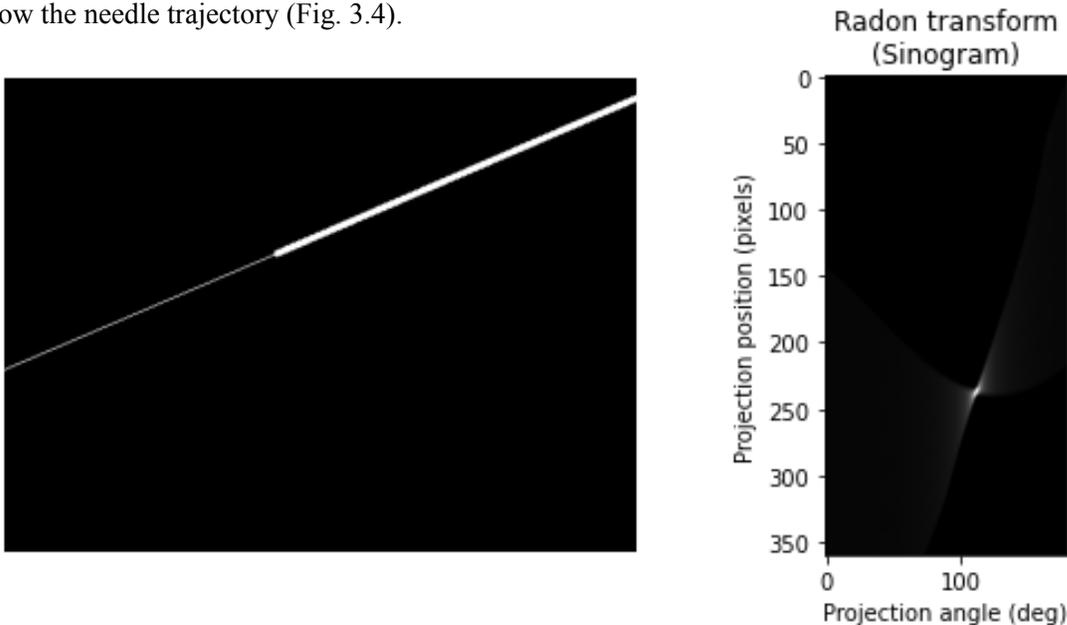


Fig 3.4 Left: Needle (thick white line) and its trajectory (thin white extension line); Right: Sinogram of needle in left image

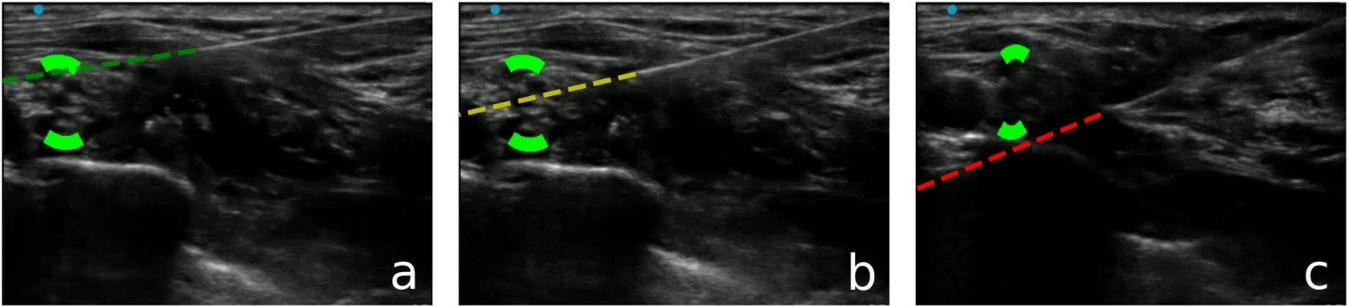


Fig. 3.5. Needle trajectory guidance towards target

As depicted in Fig. 3.5, when the needle trajectory intersects with the target anesthesia region, it is colored green. When it is between the anesthesia targets, it is colored yellow and is colored red if it misses the anesthesia targets. Supplementary video 3.1 demonstrates the needle guidance towards its target.

### 3.2.5 Evaluation

The needle trajectory was evaluated by comparing the ground truth and prediction for trajectory angle, trajectory's distance from center of the image, and needle tip coordinates as depicted in Fig. 3.6 and average errors over all frames in a video were reported.

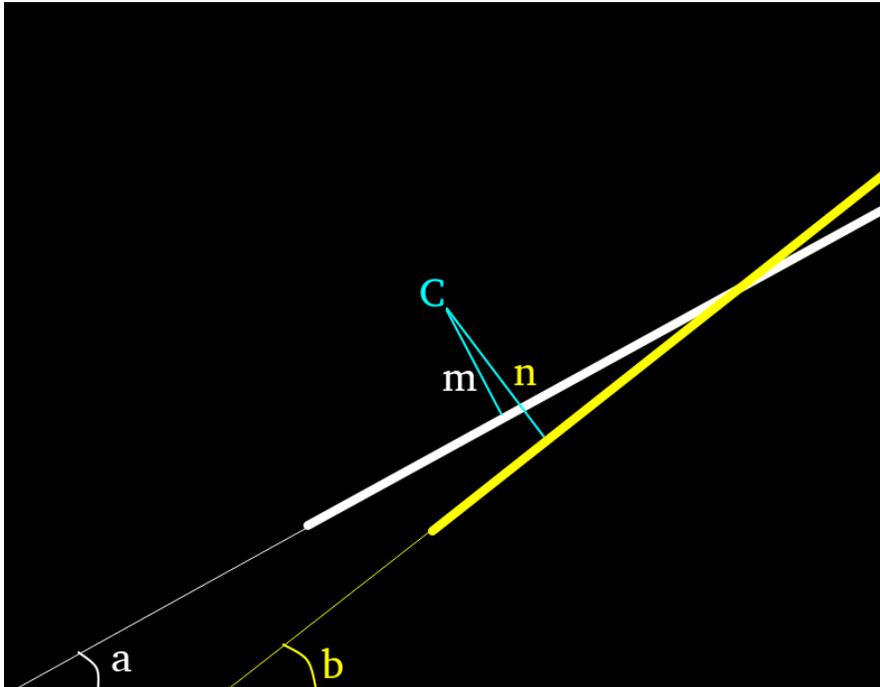


Fig. 3.6 Needle Trajectory Evaluation parameters

White thick line: Needle Ground Truth

Yellow thick line: Model's needle mask prediction

White thin line: Needle extended trajectory in ground truth

Yellow thin line: Needle extended trajectory from model mask prediction

C: Image center

a: ground truth trajectory angle

b: prediction trajectory angle

m: ground truth trajectory distance from image center

n: prediction trajectory distance from image center

### 3.3 Results

After fine-tuning on 937 frames from 5 random videos, we tested the SCBP and needle on 124 frames in the sixth video. The average dice coefficient of SCBP segmentation was computed and found to be 0.78 and the average F-score at 50% IOU was found to be 0.96. The needle trajectory was evaluated by

comparing the ground truth and prediction for trajectory angle, trajectory's distance from center of the image, and needle tip coordinates.

The needle trajectory's average angle error was  $2^\circ$ , average error in trajectory's distance from center of the image was 10 pixels (2 mm) and the average error in needle tip was 19 pixels (3.8 mm) which is within acceptable range of 5 mm as per experienced anesthesiologists.

The average run-rate was measured to be more than 30 fps using the NVIDIA RTX 3090 GPU.

### **3.4 Discussion and Conclusion**

In this chapter, a method for real-time detection of nerves together with a needle and its trajectory in ultrasound videos is presented. The needle target is precisely defined and the extrapolated needle trajectory enables real-time guidance towards it. We also publicly release needle annotations for further research.

During our experiments, we found that getting the needle trajectory based on segmentation alone led to erroneous results due to under and over segmentation of the needles which in turn is due to few pixels occupied by needles. We optimized the segmentation decision threshold to allow some under-segmentation but avoid over-segmentation completely because Radon transformation can compute the needle trajectories even with small needle parts. This is ideal for challenging cases when only a part of the needle is in the field-of-view and the model generates fragmented segmentation predictions.

#### **Limitations**

The present method has some limitations. The model is sensitive to linear artifacts, high-intensity bones and borders of lungs (pleura) and soft tissue. The model needs visibility of a few millimeters of needle to begin detecting it and may not work during the initial insertion i.e. skin puncturing. The short length visual of the needle when it is inserted is more sensitive to background noise. Reverberation artifacts from the needle may decrease the quality of the results

## **Further Research**

While the results of this research are encouraging, further research can improve upon these results. Algorithms utilizing spatio-temporal information should be investigated to exploit the temporal information present in the video data. This can help in cases when needles go out of field-of-view. For testing the generalizability of the model on new US machine data, we need to build datasets from new machines and test with and without fine-tuning. The invasive nature of needle makes this harder to get. Our dataset contains needle trajectories between  $30^\circ$  and  $60^\circ$ . More variance in trajectory angles should be tested. Needle localization is challenging at steep angles and deep insertions.

This chapter focussed on ultrasound image interpretation, localization of the needle target and guidance of the needle towards the needle target. Further research is warranted on autonomous delivery of the anesthetic, while trying to avoid puncturing the lungs, poisoning the arteries and injuring the nerves.

## *Chapter 4*

### **Conclusions**

In this thesis, we examined the computer vision related parts of developing an autonomous robot capable of UGRA i.e. ultrasound image interpretation, definition and localization of the needle target, localization of needle and its tip and guidance of the needle towards the needle target.

First, we collected 227 ultrasound videos data from different patients with the help of All India Institute of Medical Sciences, New Delhi. We created an object tracking based video annotation tool using OpenCV for the anesthesiologists which helped to automate 97%-98% of the manual annotation work. This led to creation of a large dataset comprising of 41,000 frames with manually vetted annotations of brachial plexus.

From a wide range of available off-the-shelf segmentation models, we shortlisted two models - DeepLabv3 and UNet, and their enhancements - DeepLabv3+ and UNet++ based on established usage in medical image segmentation literature, real-time processing capability and low memory usage. We tested the performance of each of these models with 5-fold cross-validation. For testing the generalizability of the model, we tested it on data from two separate scanners - a traditional cart based ultrasound machine with superior image quality and a handheld portable scanner that can be plugged into smartphones and run with an app with a compromise of a lower image quality. We achieved an acceptable dice score of brachial plexus segmentation with all the machines.

Next we focus on the needle guidance towards needle targets. First, we created a needle annotation tool using OpenCV focussed to create needle annotations of 6 ultrasound SCBP videos comprising of 1,061 frames. Then neural network training on this data was performed in two steps. A natural RGB images pre-trained UNet++ network was first fine-tuned on a large dataset of 196 ultrasound videos wherein the

network learned the ultrasound domain and to segment SCBP and ISC nerves as two output channels. In the next step, a second fine-tuning on five needle videos was done to adapt the second output channel of ISC to predict needle mask. After training on five random videos we tested the needle in the sixth video.

We define automated anesthesia needle targets derived from CNN nerve segmentation predictions by fitting an ellipse to the nerve contours. The problem of the partial visibility of the needle was overcome by trajectory extrapolation using the visible part. Radon transform was applied on the network's needle mask prediction to obtain a sinogram. Assuming that the sinogram's maximum value corresponds to the needle trajectory's angle and that the lowest point in the needle mask is the needle tip, a line is drawn from the needle tip to show the needle trajectory. Finally, the intersection of extrapolated trajectory with the needle target guides the needle navigation for drug delivery.

We have publicly released both the datasets discussed above for further research by the research community. To the best of our knowledge, these are the largest publicly released ultrasound video datasets created from in-vivo human subjects containing brachial plexus annotations and needle annotations.

#### **4.1 Future Work**

Algorithms utilizing spatio-temporal information should be investigated to exploit the temporal information. This can help in cases when needles get blurred or when they go out of field-of-view. For generalizability on new US machine data, testing on other machines data is required. The study can be extended to other nerve blocks.

This work focussed on interpreting ultrasound images, locating the needle target, and guiding the needle to the target. Further research is needed to develop an autonomous robotic system that can deliver anesthetics without puncturing the lungs, poisoning the arteries, or injuring the nerves.

## **Related Publications**

Abhishek Tyagi, Abhay Tyagi, Manpreet Kaur, & Jayanthi Sivaswamy. Towards Autonomous Robotic Ultrasound Guided Regional Anesthesia using Real Time Needle Localization and Trajectory Extrapolation. *Robot Assisted Medical Imaging workshop at IEEE International Conference on Robotics and Automation Workshop (ICRA) 2023*

Abhishek Tyagi, Abhay Tyagi, Manpreet Kaur , Jayanthi Sivaswamy , Richa Aggarwal , Kapil Dev Soni, & Anjan Trikha. Automated Real Time Delineation of Supraclavicular Brachial Plexus in Neck Ultrasonography Videos: A Deep Learning Approach. *Submitted to Nature Scientific Reports December 2022*

## Bibliography

- [1] Joshi, G. Gandhi, K. Shah, N. Gadsden, J. & Corman, S.L. Peripheral nerve blocks in the management of postoperative pain: challenges and opportunities. *J Clin Anesth.* **35**:524-529 (2016).
- [2] Chan, E.Y. Fransen, M. Parker. D.A. Assam, P.N. & Chua, N. Femoral nerve blocks for acute postoperative pain after knee replacement surgery. *Cochrane Database Syst Rev.* **(5)**:CD009941. (2014)
- [3] Chen, X. X., et al. Ultrasound-Guided Regional Anesthesia Simulation Training: A Systematic Review. *Regional anesthesia and pain medicine*, 42(6), 741–750 (2017).
- [4] Bowness, J. El-Boghdadly, K. Burckett-St Laurent D. Artificial intelligence for image interpretation in ultrasound-guided regional anaesthesia. *Anaesthesia* **76**:602–7 (2021).
- [5] Wegener, J.T. et al. Value of an electronic tutorial for image interpretation in ultrasound-guided regional anesthesia. *Reg Anesth Pain Med.* **38(1)**:44-49 (2013).
- [6] Nix, C. M., et al. A scoping review of the evidence for teaching ultrasound-guided regional anesthesia. *Regional anesthesia and pain medicine*, **38(6)**, 471–480. (2013).
- [7] Neal, J.M. et al. Upper Extremity Regional Anesthesia. *Regional Anesthesia & Pain Medicine* **34**:134–70. (2009)
- [8] Lanz, E., Theiss, D. & Jankovic, D. The Extent of Blockade Following Various Techniques of Brachial Plexus Block. In: Van Kleef, J.W., Burm, A.G.L., Spierdijk, J. (eds) Current Concepts in Regional Anaesthesia. Developments in Critical Care Medicine and Anaesthesiology, vol 7. Springer, Dordrecht. (1984)
- [9] Yang, W.T. Chui, P.T. Metreweli, C. Anatomy of the normal brachial plexus revealed by sonography and the role of sonographic guidance in anesthesia of the brachial plexus. *American Journal of Roentgenology* **171**:1631–6 (1998)
- [10] Litjens, G. et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* **(42)** 60-88 (2017)
- [11] Shen, D. Wu, G. Suk, H. Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng.* **(19)** 221–248 (2017)
- [12] Ker, J. Wang, L. Rao, J. & Lim T. Deep Learning Applications in Medical Image Analysis. *IEEE Access* **(6)** 9375-9389 (2018)
- [13] <https://www.kaggle.com/c/ultrasound-nerve-segmentation/>. Accessed on July 18, 2023
- [14] Baby, M. & Jereesh, A.S. Automatic nerve segmentation of ultrasound images. In *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)* (2017)
- [15] Kakade, A. & Dumbali, J. Identification of nerve in ultrasound images using u-net architecture. In *International Conference on Communication information and Computing Technology (ICCICT)* (2018)

- [16] Wu, H., Liu, J., Wang, W., Wen, Z., & Qin, J. Region-aware Global Context Modeling for Automatic Nerve Segmentation from Ultrasound Images. In *AAAI Conference on Artificial Intelligence (AAAI)* (2021)
- [17] Ding, Y. et al. MallesNet: A multi-object assistance based network for brachial plexus segmentation in ultrasound images. *Med. Image Anal.* **80** (2022)
- [18] Smistad, E. Johansen, K.F. Iversen, D.H. & Reinertsen, I. Highlighting nerves and blood vessels for ultrasound-guided axillary nerve block procedures using neural networks. *Journal of Medical Imaging.* **5(4)**:044004 (2018)
- [19] Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2015)
- [20] Kapoor, A. & Singh, T. A brief review: Speckle reducing filtering for ultrasound images. In *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)* (2017)
- [21] Luo, W., et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *J Med Internet Res* **18(12)**:e323 (2016).
- [22] <https://opencv.org/> .Accessed on July 18, 2023
- [23] <https://github.com/Regional-US/> Accessed on July 18, 2023
- [24] Lukezic, A. Vojir, T. Zajc, L.C. Matas, J. & Kristan, M. Discriminative Correlation Filter with Channel and Spatial Reliability. In *IEEE Conference on Computer Vision and Pattern Recognition. (CVPR)* (2017)
- [25] Henriques, J.F. Caseiro, R. Martins, P. & Batista, J. High-speed tracking with kernelized correlation filters. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* (2014).
- [26] Márquez-Neila, P. Baumela, L. & Álvarez, L. A Morphological Approach to Curvature-based Evolution of Curves and Surfaces. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* (2014)
- [27] Radon, J., translated by Parks, P.C., "On the determination of functions from their integral values along certain manifolds", *IEEE Transactions on Medical Imaging*, 5 (4): 170–176 (1986)
- [28] Chen, L.C. Papandreou, G. Schroff, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. arXiv preprint [arXiv:1706.05587](https://arxiv.org/abs/1706.05587) (2017).
- [29] Chen, L.C. Zhu, Y. Papandreou, G. Schroff F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Eur. Conf. Comput. Vis. (ECCV)* (2018)

- [30] Zhou, Z. Siddiquee, M. M. R. Tajbakhsh, N. & Liang, J. “UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imag.* (2019)
- [31] Liu, X., Song, L., Liu, S., & Zhang, Y. A review of deep-learning-based medical image segmentation methods. *Sustainability* **13(3)**, 1224 (2021)
- [32] Wang, R. et al. Medical image segmentation using deep learning: A survey. *IET Image Process.* **16**, 1243–1267 (2022)
- [33] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
- [34] Holschneider, M., Kronland-Martinet, R., Morlet, J., Tchamitchian, P.: A real-time algorithm for signal analysis with the help of the wavelet transform. Chapter In: *Wavelets:Time-Frequency Methods and Phase Space*. 289-297. (Springer, 1989)
- [35] Grauman, K. & Darrell, T. The pyramid match kernel: Discriminative classification with sets of image features. In *International Conference on Computer Vision (ICCV)* (2005).
- [36] Lazebnik, S. Schmid, C. & Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2006)
- [37] Chen, L.C. Papandreou, G. Kokkinos, I. Murphy, K. & Yuille, A. L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. In *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* (2016)
- [38] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* (NeurIPS) (2019)
- [39] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In: *IEEE Conference on Computer Vision and Pattern Recognition*.(CVPR) (2009)
- [40] Kingma, Diederik P. and Ba, Jimmy. Adam: A Method for Stochastic Optimization - In *International Conference on Learning Representations (ICLR)* (2015)
- [41] Le, M-P.T. et al. Comparison of four handheld point-of-care ultrasound devices by expert users. *The Ultrasound Journal* **14** (2022)
- [42] Newhouse, S.M. Effing, T.W. Dougherty, B.D. D’Costa, J.A. & Rose, A.R. Is bigger really better? Comparison of ultraportable handheld ultrasound with standard point-of-care ultrasound for evaluating safe site identification and image quality prior to pleurocentesis. *Respiration.* **99(4)**:325–32 (2020)

- [43] Padilla, R. Netto, S.L. da Silva, E.A.B. A Survey on Performance Metrics for Object-Detection Algorithms," In *International Conference on Systems, Signals and Image Processing (IWSSIP)* (2020)
- [44] Lu, J. et al. Transfer learning using computational intelligence: A survey. *Knowledge-based Systems*; **80** 14–23 (2015)
- [45] Malik, H. et al. A Comparison of Transfer Learning Performance Versus Health Experts in Disease Diagnosis From Medical Imaging. In *IEEE Access* (**8**) 139367-139386 (2020)
- [46] G. Reusz P. Sarkany, J. Gal, A. Csomos, Needle-related ultrasound artifacts and their importance in anaesthetic practice, *BJA: British Journal of Anaesthesia*, 112(**5**), 794–802 (2014)
- [47] Chaurasia Manoj, Surgery without anaesthesia in Bihar hospital: Pinned down by four people and shrieking, women operated on. *Times of India*, Nov 17, 2022
- [48] Viderman, D. et al "Artificial intelligence in ultrasound-guided regional anesthesia: A scoping review." *Frontiers in Medicine* (2022)