UCLA
UCLA Electronic Theses and Dissertations

Title

Deep Learning-Based Automatic Pipeline for 3D Needle Localization on Intra-Procedural 3D MRI

Permalink

https://escholarship.org/uc/item/17j136v9

Author

Zhou, Wenqi

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Deep Learning-Based Automatic Pipeline for 3D Needle Localization on

Intra-Procedural 3D MRI

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Bioengineering

by

Wenqi Zhou

2023

ABSTRACT OF THE THESIS

Deep Learning-Based Automatic Pipeline for 3D Needle Localization on

Intra-Procedural 3D MRI

by

Wenqi Zhou

Master of Science in Bioengineering

University of California, Los Angeles, 2023

Professor Holden H. Wu, Chair

The distinct advantages of magnetic resonance imaging (MRI), including exquisite soft tissue contrast, diverse contrast mechanisms, and tomographic imaging with adjustable plane orientation, have made it a promising technology for guiding and directing interventions and surgeries. Nonetheless, MRI-guided percutaneous interventions encounter challenges in obtaining accurate, real-time 3D needle localization due to the intricate structure of biological tissue and the variability of needle features on *in vivo* MR images. This thesis aims to develop and assess a deep learning-based automatic pipeline for rapid 3D needle localization using intra-procedural 3D MRI.

Firstly, the pipeline incorporated Shifted-window UNEt TRansformers (Swin UNETR)

for 3D needle feature segmentation on intra-procedural 3D MRI. Next, a post-processing method was developed to determine and extract the 2D reformatted image plane that passes through the main axis of the segmented 3D needle feature. Lastly, a 2D Swin Transformer network was adapted and trained for fine segmentation of needle features on reformatted 2D image planes. The 3D needle location was calculated based on the 2D coordinates of the needle feature tip and entry point on the 2D reformatted images.

The pipeline was evaluated using *in vivo* 3D MR images acquired during MRI-guided interventional experiments in pre-clinical pig models. The automatic pipeline achieved real-time and accurate 3D needle localization, with a needle tip and axis localization accuracy comparable to human intra-reader variations and a processing time of about 6 seconds from start to end. This pipeline can offer real-time visual assistance for physicians during percutaneous procedures, expediting the interventional workflow and holding the potential for enabling robotic-assisted MRI-guided interventions.

The thesis of Wenqi Zhou is approved.

Dan Ruan

Kyunghyun Sung

Holden H. Wu

University of California, Los Angeles

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1 Significance

Over the past 20 years, advances in medical imaging have significantly expanded the field of image-guided interventions[1]. Image-guided interventions play a key role in enabling new minimally invasive procedures with their markedly less invasive nature[2]. Traditional image-guided interventions use imaging modalities such as computed tomography (CT) and ultrasound (US)[3]. However, US offers relatively low-quality images of tissues and needles[4], limiting its ability to visualize deep or small lesions[5]. Meanwhile, CT may fail in visualizing certain lesions, such as hepatocellular carcinoma (HCC)[6]. On the other hand, MRI excels in visualizing various cancerous lesions in organs like the liver or kidney, detecting smaller tumors that CT and US cannot[7]. Furthermore, MRI eliminates ionizing radiation exposure for both physicians and patients[8] and allows for 2D or 3D imaging in arbitrary orientations.

Despite MRI's distinct advantages, performing MRI-guided percutaneous interventions in the abdomen remains challenging due to difficulties in real-time needle localization[9]. Manual interventional needle localization on 3D images is time-consuming, as it involves manually inspecting multiple slices of the 3D volume to identify the needle tip and axis, potentially prolonging interventional procedures[10]. Previous research has investigated 3D deep learning-based methods for automatic interventional needle localization in 3D MR images[11]. However, training 3D deep learning networks generally requires a large-scale dataset of 3D MR images to achieve reliable

performance due to the complexity of biological tissue structures and the variability in needle features' appearance on *in vivo* MR images. However, a large 3D MRI dataset may not be available for certain applications[12].

This study aims to expand the technical capability of MRI to guide percutaneous interventions by developing a pipeline that automatically localizes interventional needles on 3D images in real time. Utilizing transformer-based deep learning networks and a coarse-to-fine pipeline structure, the pipeline was trained using a relatively small set of *in vivo* 3D MR images and demonstrates high accuracy in real-time needle feature tip and axis localization.

1.2 MRI-Guided Percutaneous Interventions

1.2.1 Overview of MRI-Guided Needle Placement

In MRI-guided percutaneous interventions, needles are inserted into the target lesion under MRI guidance to perform tissue biopsy or deliver therapeutic energy to the lesion[13]. MRI offers excellent soft tissue contrast, enabling the visualization of various types of cancerous lesions and critical structures that must be avoided during interventions. Additionally, MRI provides needle visualization based on passive signal void features caused by needle-induced magnetic susceptibility effects, guiding needle manipulation during procedures. The capacity to obtain image slices in arbitrary orientations without ionizing radiation has positioned MRI as an emerging intra-procedural imaging modality[14].

In typical MRI-guided interventions, physicians have established the step-and-shoot workflow to adapt the MRI advantages for percutaneous interventions[1]. The workflow includes the planning, insertion, and confirmation stage[13] as illustrated in **Figure 1-1**:
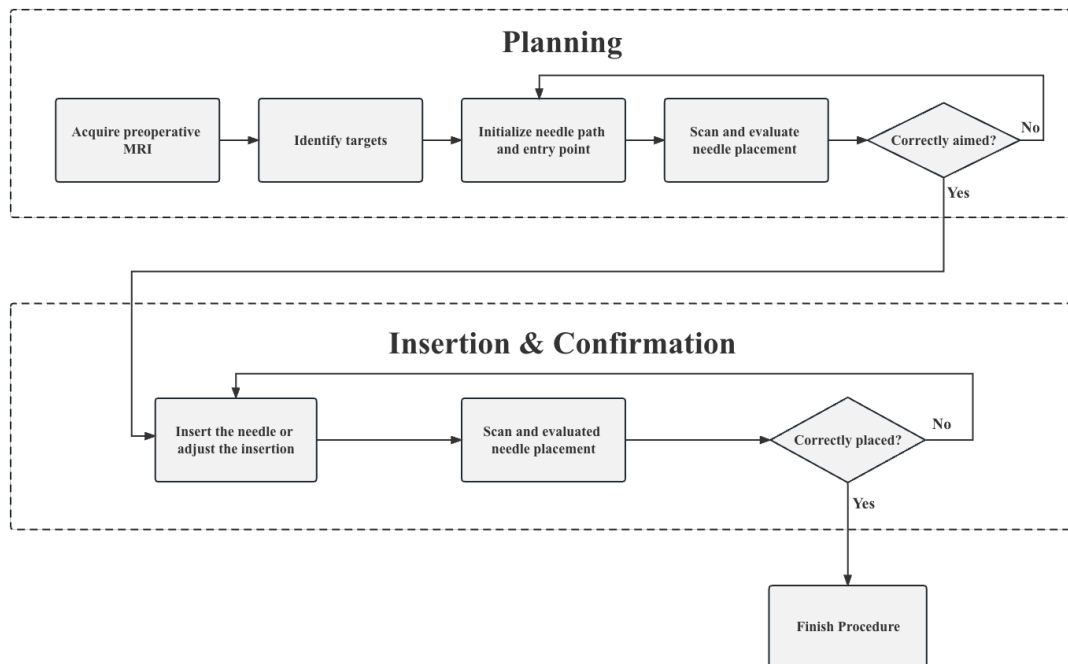
Figure 1-1 The step-and-shot workflow for MRI-guided percutaneous interventions.

During the step-and-shoot workflow of the MRI-guided intervention procedure, the patient or experimental subject is placed on the imaging table and may receive local or general anesthesia. Preoperative MR images of the region of interest are acquired for planning purposes. The physician uses the preoperative MR images to identify suspicious target regions and initialize a suited entry point on the patient's skin. The needle trajectory is then chosen such that impenetrable structures such as bones and critical structures such as larger blood vessels are avoided[15].

Next, the physician inserts the needle according to the chosen path and scans the region of interest to visualize the position of the needle relative to the patient's anatomy. The physician then adjusts the needle trajectory in a step-and-shot procedure, with intermediate confirmation scans and angular adjustments. The intermediate confirmation scans were done with the patients in the MR scanner bore, while the needle insertion and adjustment require the patient table to be moved out from the MRI scanner bore. This 'in/scan–out/adjust' technique[16] inhibits physicians from interactively manipulating the needles while visually monitoring the needle location relative to the target based on direct image feedback. Typically, a repeated in/scan–out/adjust process is required until the needle reaches the intended location. The number of required iterations largely depends on the physicians' skill in localizing the needle in 3D and performing desired needle adjustments and insertions[17,18], which may lead to prolonged procedure time.

The iterative step-and-shoot process of manual adjustments to needle placement and subsequent validation has three direct negative consequences: (a) lack of direct image feedback during the manipulation of the needle; (b) increased tissue damage due to repeated incorrect needle insertions; (c) greater patient discomfort and higher costs as a result of extended procedure duration, emphasizing the need to enhance procedural efficiency and accuracy/precision by developing systems that assist in percutaneous needle placement[12].

1.2.2 Clinical Applications of MRI-Guided Percutaneous Interventions

MRI-Guided Percutaneous Biopsy

MRI-guided biopsy is becoming increasingly important for the diagnosis and treatment of various abdominal lesions. Traditional imaging methods, such as US, have limitations in detecting more than half of small focal liver lesions[14]. In CT-guided biopsy, the iodine-based contrast agent may be required in some cases, making it unsuitable for patients who are allergic to iodine[20].

MRI facilitates visualization and access to abdominal targets by improving biopsy accuracy with excellent soft-tissue contrast[19]. Recent studies have demonstrated the feasibility of MRI-guided biopsy in various abdominal organs, including the liver, gallbladder, pancreas, spleen, and kidney, with targeting accuracy rates ranging from 89% to 100%[14]. The median duration of MRI-guided biopsy procedures is less than 60 minutes, which is comparable to that of CT-guided biopsy. However, MRI-guided biopsy also has challenges, including respiratory motion and longer distances between the skin and targeted lesions, particularly in the liver. Overall, MRI-guided biopsy is a promising tool for accurate diagnosis and treatment of abdominal lesions, although further research is needed to optimize the technique and overcome its challenges.

MRI-Guided Percutaneous Thermal Therapy

MRI-guided ablation methods, including percutaneous radiofrequency ablation (RFA), and microwave ablation (MWA), laser-induced thermotherapy (LITT), have been used in the treatment of cancer and other disease.[21,22,23]. MRI-guided percutaneous

radiofrequency ablation (RFA) is a minimally invasive therapy that uses an electromagnetic field to induce thermal injury[22]. CT or US is commonly used to guide RFA. On the other hand, there is ongoing development of MRI-guided RFA[24]. Multiple reports have been published on the feasibility and safety of RFA of liver lesions in a wide-bore MRI scanner[14]. T_1-weighted (T1w) imaging is generally used to visualize thermal changes in tissue during MRI-guided percutaneous RFA. T1w MRI after RFA can be used to measure the size of the ablation zone and guide the RFA procedure[14,25]. However, MRI-guided RFA requires switching between imaging and the application of therapeutic radiofrequency energy due to radiofrequency emission interference with imaging. MRI-guided percutaneous microwave ablation (MWA) features a higher temperature, faster ablation time, larger ablation zones, and lower susceptibility to heat sink than radiofrequency ablation (RFA). Therefore, MWA may be preferable to RFA, especially for tumors ≥ 3 cm in diameter or those located close to large vessels, regardless of size[26]. MRI-guided percutaneous laser-induced interstitial thermotherapy (LITT) is currently performed using implantable percutaneous catheter systems in local anesthesia and with outpatient management[27]. In LITT, laser energy is transmitted to the target volume via an optical fiber and is converted into thermal energy, which causes tissue damage. The extent of thermal damage is controlled through the use of real-time MR-thermography guidance. The advancement of magnetic resonance imaging thermometry in recent years has enabled precise monitoring and delivery of thermal energy to predetermined targets, thereby reviving LITT as a powerful and practical tool in neurosurgical procedures[28].

1.2.3 Advancements in MRI-Guided Percutaneous Interventions

A major limitation of MR-guided interventions is the restricted space in closed-bore MRI scanners. The space inside the MRI bore is typically 60-70 cm in diameter and 170 cm in length of the bore[29]. The limited space makes manual needle placement in the scanner bore very difficult. On the other hand, needle placement outside the bore (e.g., during step-and-shoot procedures) suffers from time-consuming, iterative needle positioning and does not permit simultaneous imaging and needle insertion[8].

To address the challenges of the manual insertion scenario, MRI-compatible remote-controlled systems are being developed to achieve needle insertion inside the scanner controlled by operators outside the scanner[30,31]. Wu et al. developed a 2-degree-of-freedom (DOF) needle driver and actuation box[32]. The 2-DOF needle driver is placed inside the scanner bore and driven by the actuation box positioned at the end of the table through a beaded chain transmission[32]. Such systems address the challenge of limited space within the bore and facilitate simultaneous MRI guidance while emphasizing the importance of real-time automatic MRI-guided device tracking for these applications.

1.2.4 Needle Feature Visualization on MR Images

The visualization of the needle relative to surrounding tissues on MRI is crucial for the safety and success of percutaneous biopsies. Typically, MRI-compatible needles used in these procedures are metallic and have a magnetic susceptibility different from

water[33]. The difference in magnetic susceptibility between the needle and the surrounding tissue results in magnetic field perturbation and leads to MR signal dephasing[34]. Passive signal void features on MR images, caused by the susceptibility effects induced by the needle, are used to indicate the location of the needle.

Multiple factors can affect the appearance of the needle feature on MRI. The higher the MR field strength, the larger the apparent width of the needles[8]. The frequency encoding direction also affects the size of the needle feature. When a needle is present in the tissue, it distorts the local magnetic field, which results in image artifacts in the frequency-encoding direction. This effect is more pronounced when the frequency encoding is perpendicular to the needle, causing the artifact to appear wider. If the frequency encoding is parallel to the needle, the artifact is less noticeable [35,36].

The apparent needle diameter is also related to the relative needle orientation to the main magnetic field (B_0). As **Figure 1-2** shows, when the needle moves closer to the axis of the main magnetic field, the diameter of the needle feature decreases significantly due to a reduction in image artifact. The distortion in the magnetic field that causes artifacts are most prominent at the points where the field enters and exits objects with different magnetic susceptibilities, such as the needle and the surrounding tissue. When the needle is aligned parallel to the main magnetic field, the distortion of the magnetic field (and therefore image artifact) is mainly present at the tip of the needle, with less distortion along the needle shaft[37].

8
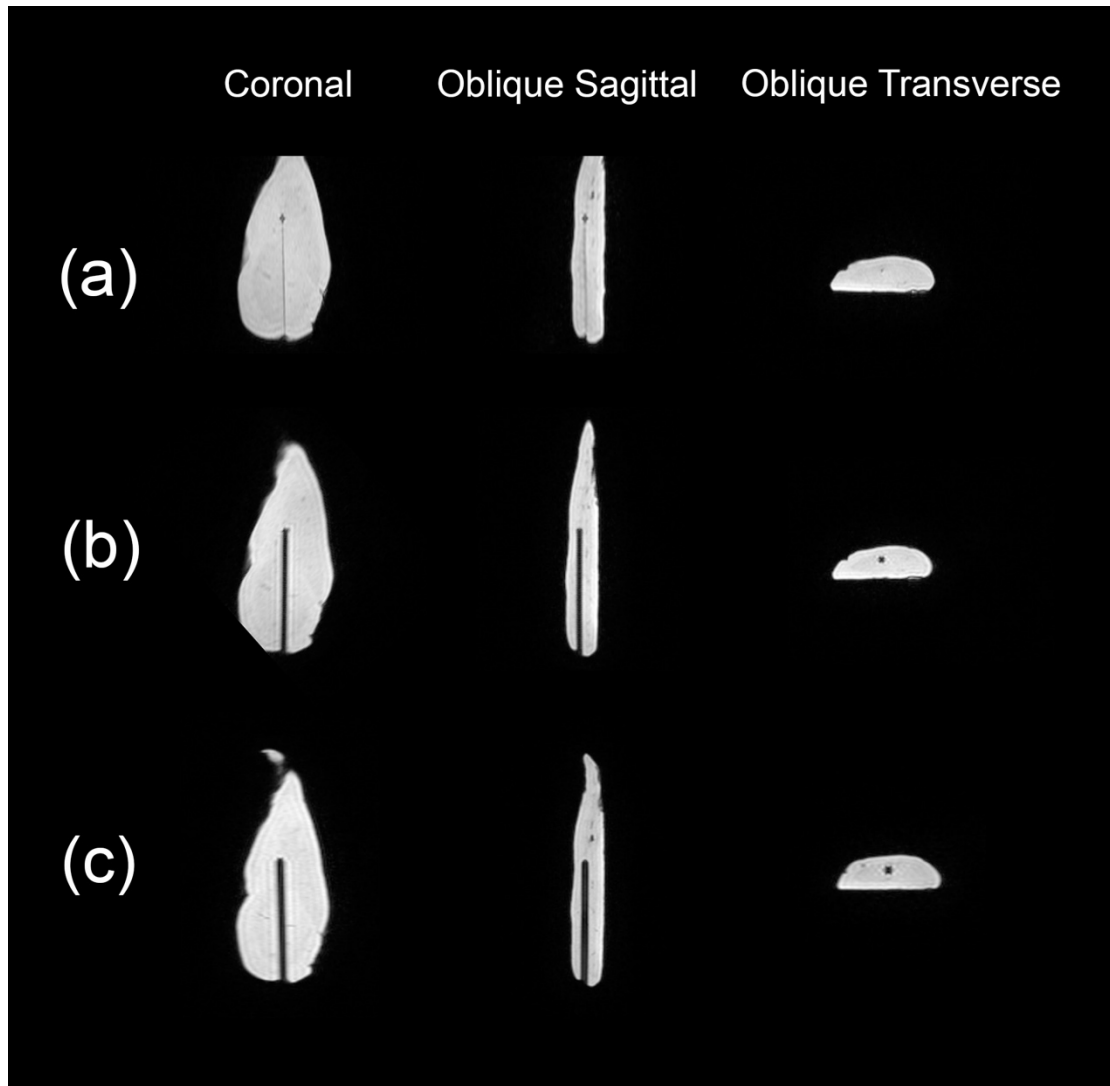
Figure 1-2 3D T1w gradient-echo MR images of a 22 gauge, 15 cm needle (Mreye®) with different needle axis orientations with respect to the B_0 field. (a)Needle axis parallel with the B_0 field. (b) The angle between the needle axis and the B_0 field is approximately 45 degrees. (c) Needle axis is perpendicular to the B_0 field.

1.2.5 Automatic Needle Localization Techniques

Accurate and efficient needle placement is crucial for the success of percutaneous interventional procedures, and MRI guidance allows physicians to visualize the needle's feature tip and trajectory relative to the target[11]. Recent studies have leveraged deep neural networks for needle feature segmentation and localization in MRI.

For instance, Li et al. used a 2D Mask Region-based Convolutional Neural Network (R-CNN) to automatically segment needle features in real-time MRI datasets[38]. The algorithm achieved consistent needle localization performance, with an accuracy of around 1 mm in needle tip detection and around 1° in needle trajectory angle. However, a major limitation of this study was that the oblique 2D slice containing the needle feature needed to be located manually, and the output of the network was limited to 2D segmentation masks and 2D needle coordinates in the slice.

To overcome this limitation, neural networks that can perform needle localization on 3D MR images are required. Mehrtash et al. adopted a deep 3D fully convolutional neural network (CNN) for needle segmentation on 3D MR images, achieving an average needle localization accuracy of 2.80 mm in needle tip detection and 0.98° in needle trajectory angle[11]. However, one major drawback of this approach is that the large 3D dataset needed for training a fully 3D CNN is not always available for the intended application.

1.3 Advancements in Image Segmentation with Transformer Architectures

1.3.1 Vision Transformer

In recent years, transformers have gained considerable popularity due to their exceptional ability to capture long-range global information[39]. Transformers were first proposed for machine translation and have since become the state-of-the-art method in many natural language processing (NLP) tasks. The attention mechanism of the

transformer allows the model to capture the long-term dependencies of sentences. Inspired by its success in NLP, researchers further developed Vision Transformer (ViT) networks for computer vision tasks. In ViT, an image was split into many patches and the image patches are then treated as words in a sentence. By exploring global interactions between different patches, the ViT learns to direct attention to crucial image regions[40].

1.3.2 2D Swin Transformer

To tailor the transformer more effectively for computer vision tasks, researchers developed the Shifted Window (Swin) Transformer. Compared to the ViT, the Swin Transformer has three distinctive characteristics. First, self-attention is computed only within local windows, and the windows have different sizes to capture visual elements in various scales. Second, the number of patches in each window is fixed, and thus the complexity becomes linear to the image size. Third, the windows are shifted between each layer, thus providing connections among each window[40]. By contrast, ViT only produces feature maps of a single low resolution and has quadratic computation complexity with respect to input image size due to the computation of self-attention globally.

1.3.3 3D Swin UNETR

Building on the success of the Swin Transformer, researchers introduced the Swin UNEt TRansformers (Swin UNETR) to further enhance the capabilities of the

transformer architecture for medical imaging tasks[41,42]. The Swin UNETR model is characterized by a U-shaped network architecture[43], employing a Swin transformer as the encoder and a CNN-based decoder linked to the encoder via skip connections at varying resolutions. The encoder within the Swin UNETR is uniquely designed to process 3D input patches directly and can be fine-tuned using self-supervised pre-training tasks, making it highly suitable for 3D medical image segmentation. Hatamizadeh at el. demonstrated the efficacy of Swin UNETR in 3D brain tumor segmentation during the BraTS 2021 segmentation challenge, achieving top-tier performance[41]. Moreover, in the PARSE challenge 2022, Maurya et al. conducted a comparative analysis between the U-Net and Swin UNETR for pulmonary artery segmentation, with Swin UNETR outperforming the U-Net[44].

1.4 Specific Aims

This thesis aims to build and evaluate a novel coarse-to-fine deep learning-based pipeline for 3D needle feature segmentation and needle localization using limited intra-procedural 3D *in vivo* MR images collected during MRI-guided interventions. The pipeline combines an efficient 3D deep learning network responsible for initial 3D needle feature segmentation, and a 2D deep learning network that performs fine segmentation of the needle feature on the 2D slice containing needle feature.

The purpose of this work is to automatically perform 3D needle localization in real-time, providing a visual aid for physicians during percutaneous interventions and

12

accelerating the interventional procedure by reducing the time spent on manual needle localization. The coarse-to-fine nature of the pipeline and the transformer-based neural network design enables the pipeline to use a limited number of 3D images for training while achieving accurate needle localization.

1.4.1 Aim 1: To Develop and Test Neural Networks for 3D and 2D Needle Feature Segmentation on Intra-Procedural MRI.

This study adapted 3D Swin UNETR for initial 3D needle feature segmentation on intra-procedural MRI collected during MRI-guided *in vivo* pig liver interventions. The 3D model was trained and evaluated on 3D T1w Volumetric Interpolated Breath-hold Examination (VIBE) gradient-echo (GRE) MR images. As a result, the Swin UNETR model generated a 3D volumetric representation of the needle feature, enabling the localization of the oblique 2D image plane containing the needle feature.

Furthermore, this study utilized a 2D Swin Transformer network for fine segmentation of the 2D needle feature on the realigned image plane containing the needle feature to improve the accuracy of the needle tip and axis localization. The 2D Swin Transformer network was first pre-trained on single-slice 2D GRE MR images and then fine-tuned with the reformatted 2D image slices from the 3D T1w-VIBE dataset. The resulting 2D segmentation mask was used to localize the needle tip and axis location in the 2D slice and converted to coordinates in 3D space.

1.4.2 Aim 2: To Develop and Test an Automatic Deep Learning-Based Pipeline for 3D Needle Localization on Intraprocedural MRI.

The 3D Swin UNETR and the 2D Swin Transformer network specified in Aim 1 were combined with post-processing algorithms to form an automatic end-to-end pipeline that takes the intraprocedural 3D MR image as input and produces the 3D coordinates of the needle tip and the orientation of the needle axis as output. The pipeline was evaluated in the aspect of needle localization accuracy and processing time and achieved consistent needle localization performance with real-time processing time.

Chapter 2 Methods

2.1 MRI-Guided Interventional Experiments

In an animal research committee-approved study, we performed MRI-guided needle (Cook Medical, 20 gauge, 15 cm; Invivo, 18-gauge, 10 cm) interventions in the liver of seven healthy female pigs (32-36 kg) on a 3 T scanner (Prisma, Siemens). First, multiple targets were created in the liver using RFA under CT and US guidance. The targets were confirmed under post-contrast CT. Following the creation of targets, the pigs were transported to the MRI suite for needle insertion experiments. Under 3D T1w-VIBE MR images (imaging parameters listed in **Table 2-1**), the diameter of the target feature under post-contrast MRI ranged between 10.6 mm to 16.8 mm.

During the interventional experiments, the pigs were under anesthesia, and breathing was controlled by a ventilator. The MRI-guided needle insertion experiments consisted of planning, insertion, and confirmation stages, with a semi-automatic software interface providing guidance. In the planning stage, fiducial grids were attached to the pigs' skin to identify needle entry points relative to anatomical structures, while needle trajectories were planned based on entry points and targets. In the insertion stage, 3D T1w-VIBE images were used to track needle trajectories and adjust them as needed. Finally, in the confirmation stage, 3D T1w-VIBE images were used to verify needle placement within the target. The time for each targeted needle insertion, from planning to confirmation, was approximately 30 minutes.

2.2 Datasets

2.2.1 Overview

(1) 3D T1w-VIBE Dataset: During the insertion and confirmation stages, intra-procedural 3D T1w VIBE Dixon water MR images were acquired with the parameters in **Table 2-1**. For each of the seven interventional experiments, seven images were selected, resulting in a total of 49 volumetric images.

(2) 2D GRE dataset: During the insertion stage of the experiment, single-slice golden-angle 2D radial gradient-echo images[45] with the image plane aligned with the needle axis were collected with the parameters in **Table 2-1**. In each of the seven interventional experiments, seventy images were chosen, resulting in a total of 490 images.

(3) 2D reformatted image dataset: Image plane realignment was performed on the 3D T1w-VIBE images to obtain 2D reformatted images that pass through the needle axis. The generated 2D reformatted images have a field of view of 300x300 mm^2 and an in-plane resolution of 1.56x1.56 mm^2. In each of the 3D T1w-VIBE images, one 2D reformatted image containing the needle was generated, resulting in a total of 49 images.

Table 2-1 MRI Datasets and Imaging Parameters. Intra-procedural 3D T1w-VIBE MRI and 2D GRE MRI were used to train and test the proposed algorithms. TR: repetition time. TE: echo time. FOV: field-of-view.

	3D T1w-VIBE Dataset	2D radial GRE dataset
TR/TE	3.91 ms / 1.23 ms, 2.46 ms (TE1, TE2)	3.8 ms / 1.72 ms 5.08 ms / 3 ms 6.3ms / 2.85 ms
FOV	237 x 346 x 180 mm^3	300 x 300 mm^2
Number of Slices	120	1
In-plane resolution	1.35 x 1.35 mm^2	1.56 x 1.56 mm^2
Matrix Size	176 x 256	192 x 192
Slice thickness	1.5 mm	5 mm
Flip angle	9°	9°
Parallel imaging factor	4	N/A
Acquisition time	13 s (breath held)	100 ms (free breathing)

2.2.2 Image Augmentation

The size of each dataset was relatively small. Therefore, all three datasets were expanded through data augmentation by applying the following transformations: random rotation (0°–360°), horizontal flipping, vertical flipping, translation, zooming, and adding Gaussian noise. With the augmentation, the size of each dataset grew by a factor of 15. **Table 2-2** shows the size of the datasets before and after image

augmentation.

Table 2-2 Size of datasets before and after image augmentation

	3D T1w-VIBE Dataset	2D radial GRE dataset
Size of Dataset Before Augmentation	49 3D Volumes	490 2D Slices
Size of Dataset Before Augmentation	784 3D Volumes	7840 2D Slices

2.2.3 Annotation Creation

For 3D T1w-VIBE images, the 3D needle feature segmentation references were generated by a researcher under the guidance and supervision of a radiologist. First, the needle feature in each axial slice of the 3D images was annotated. Then the slice-wise annotations were reviewed and edited in the 3D Slicer software to ensure the continuity of the 3D needle feature segmentation reference. For 2D radial GRE images and 2D reformatted images, annotation of the 2D needle feature was performed according to the same guidelines. Examples of the needle feature segmentation annotations are shown in **Figure 2-1.** To create needle feature tip and needle axis references, the needle feature tips and entry points (at the skin) in 3D T1w-VIBE Dixon water images were annotated in 3D Slicer[46] by a researcher according to previously established guidelines[38,47]. The angle and insertion depth of the needle axis were calculated based on the 3D coordinates of the needle tip and needle entry point. The researcher repeated

18

the annotations with a washout period of two weeks to assess the human intra-reader

variation. Examples of the needle tip and axis annotations are shown in **Figure 2-2**.
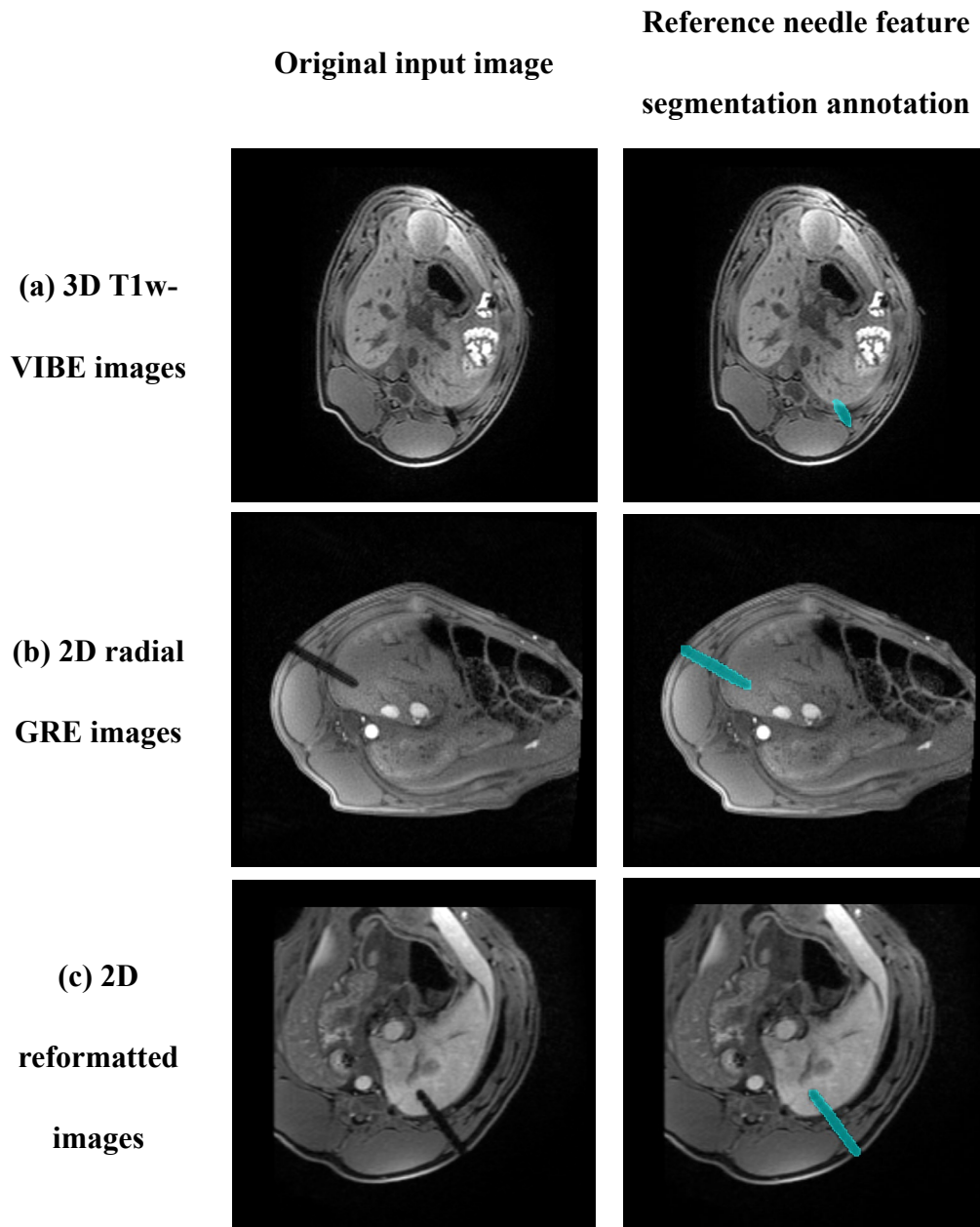
Figure 2-1 Examples of the reference needle feature segmentation annotations on 3D T1w-VIBE images, 2D radial GRE images, and 2D reformatted images.

Needle Tip Annotation
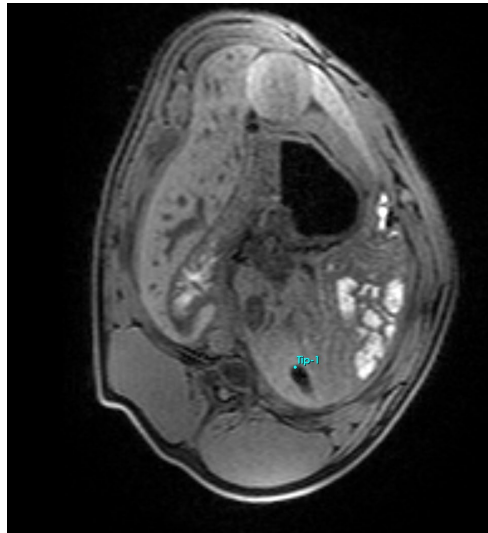
Needle Entry Point Annotation
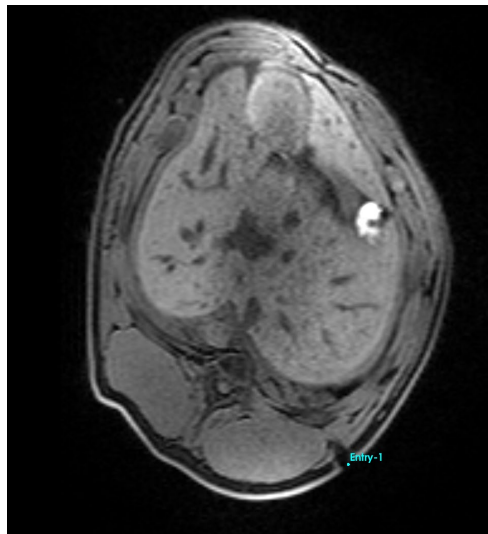
Needle Entry Point Annotation
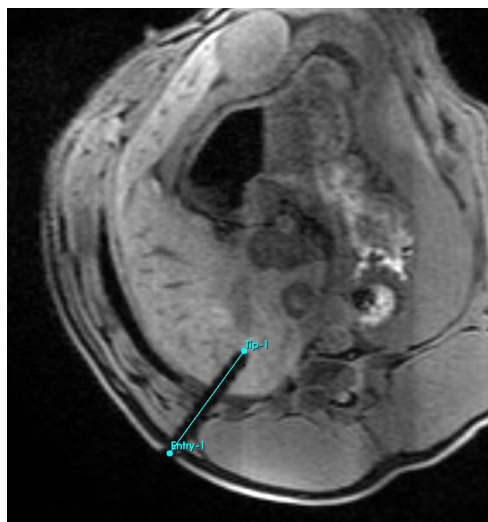
Figure 2-2 Examples of the reference needle tip and needle axis annotations on 3D T1w-VIBE images. Tip: needle tip; Entry: needle entry point.

2.3 Overview of 3D Needle Localization Pipeline

The proposed pipeline (**Figures 2-3** and **2-4**) takes 3D T1w-VIBE Dixon water images as input and localizes the needle feature tip and axis in 3D space via a fully automatic process implemented in 3D Slicer[46]. There are three main steps in the pipeline:

Step 1: An initial 3D segmentation of the needle feature was generated with the 3D Swin UNETR and post-processed with the false-positive removal module.

In the first step, we applied the 3D Swin UNETR to the 3D T1w-VIBE image to segment the needle feature and then used post-processing modules to remove false-positive segmentations. These modules determine the volume of each object in the segmentation results and remove the small ones, as true positive segmentations of the needle tend to have the largest volume. The result of Step 1 is a coarse 3D segmentation of the needle feature, with false positives removed. Any false negatives (under-segmentation) will be addressed in the following step, which involves fine segmentation.

Step 2: A 2D reformatted image plane that passes through the main axis of the needle feature was generated from the 3D image.

To perform image plane realignment, we first fitted a bounding box to the 3D needle feature segmentation and identified the main axis of the bounding box as the axis of the needle feature. Next, we carried out oblique axial image plane realignment, resulting in the generation of a 2D reformatted image plane that passes through the axis of the

needle feature. This 2D reformatted image plane will serve as the input for the third step of the pipeline.

Step 3: The 2D coordinates of the needle feature tip, entry point, and axis were calculated from the 2D network segmentation mask and converted to 3D coordinates.

We applied the 2D Swin Transformer network to the reformatted 2D image generated in Step 2 to generate a 2D segmentation mask of the 2D needle feature. Next, we employed orthogonal distance regression (ODR)[48] to localize the needle axis in the 2D image using the segmentation mask. The intersection of the detected needle axis and the 2D segmentation mask was identified as the needle feature tip and needle entry point. We then converted the 2D coordinates of the needle tip and entry point into 3D coordinates, producing the needle tip coordinates and needle axis orientation in 3D space.
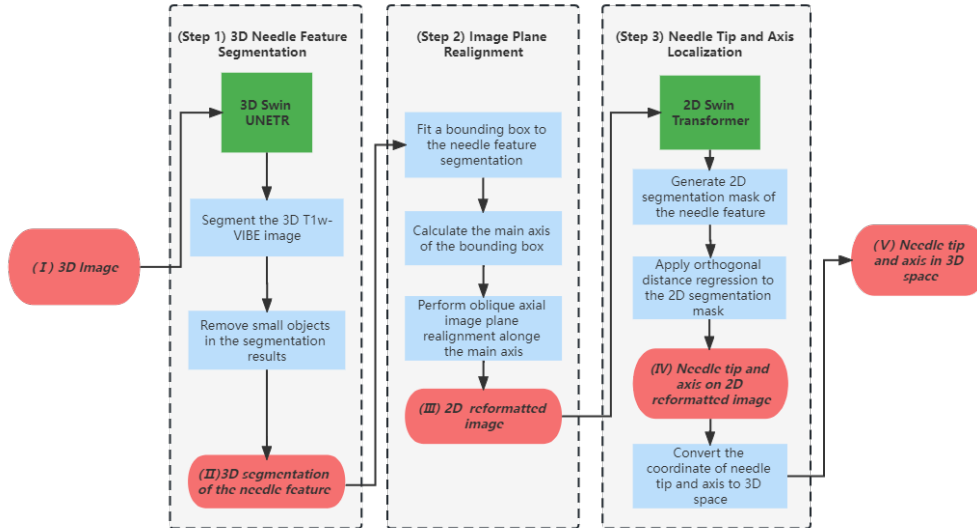
Figure 2-3 Schematic showing the deep learning-based automatic pipeline for coarse-to-fine needle localization in 3D space using intra-procedural 3D MRI.
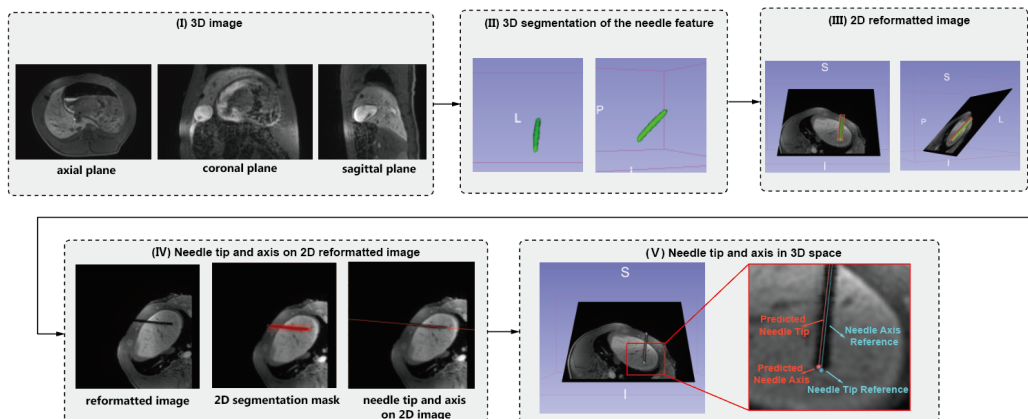
Figure 2-4 Output of each step in the proposed automatic pipeline. (I) 3D T1w-VIBE Dixon water image. (II) Deep learning-based 3D segmentation of the needle feature displayed in 3D space using subject coordinates (R/L: right/left A/P: anterior/posterior, S/I: superior/inferior). (III) The 2D reformatted image aligns with the axis of the segmented needle feature. (IV) Deep learning-based localization of the needle tip and axis on the 2D reformatted image. (V) Needle tip and axis in 3D space.

2.3 3D Swin UNETR

2.3.1 3D Swin UNETR Network Structure

We employed the 3D Swin UNETR, implemented using Keras and PyTorch

frameworks[49], as depicted in **Figure 2-5**. The input grayscale 3D MR images were zero-padded into a matrix of 256 x 256 x 128 pixels by adding zeros, and the needle signal void feature on 3D MR images was defined as the only non-background class.

In the encoder of the Swin Transformer, the patch partitioning layer created a sequence of 3D tokens, then the 3D tokens were divided into non-overlapping windows and local self-attention was computed within each region. Specifically, the partitioned windows were shifted between each layer by one-half of the window size. In between each stage, a patch merging layer was used to reduce the resolution by a factor of 2. The hierarchical representations of the encoder at different stages were used in downstream applications such as segmentation for multi-scale feature extraction[40].

The CNN-based decoders were connected to the encoder via skip connections. The extracted features were processed using a residual block with 3x3x3 convolutional layers and instance normalization[50]. The processed features were up-sampled using a deconvolutional layer and concatenated with the previous stage's features and then passed through another residual block before feeding to the next stage. For segmentation tasks, the Swin Transformer encoder output was first concatenated with the processed input features. The resulting feature map was then passed through a residual block, followed by a 1x1x1 convolutional layer with softmax activation that produced segmentation probabilities.
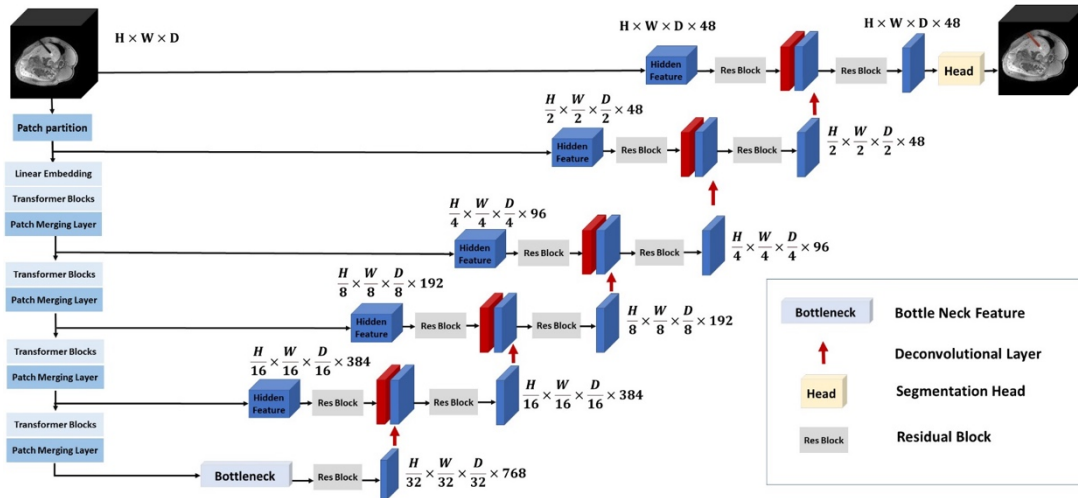
Figure 2-5 Overview of the 3D Swin UNETR architecture[42]. The Swin UNETR architecture processed 3D T1w-VIBE MRI images as input. It generated distinct patches from the input data to establish windows of a specific size for self-attention calculation. The Swin transformer's encoded feature representations were then transmitted to a CNN decoder through skip connections at various resolutions. W:255, H:256, D:128.

2.3.2 3D Swin UNETR Training and Testing

We adopted the pre-trained 3D Swin UNETR model that was pre-trained on 5,050 publicly available CT images from various body organs with self-supervised learning tasks[42]. These tasks include masked volume inpainting, 3D image rotation, and contrastive coding. The pre-training was performed by minimizing the total loss function of the pre-training tasks. To train the 3D Swin UNETR, we employed the Dice Loss[51] as the training loss metric. Additional specific training hyperparameters are shown in **Table 2-3**.

To evaluate the 3D needle feature segmentation using 3D Swin UNETR, we performed

seven-fold cross-validation (seven 3D T1w-VIBE images in each fold). All images in the training dataset underwent 15-fold image augmentation. The Dice similarity coefficient (Dice, 0 to 1) was used as a metric to assess experimental results.

Table 2-3 3D Swin UNETR model training parameters.

Batch size	**4**
Iteration	**10000**
Weight decay	**0.00001**
Learning rate	**0.0001**
Window size	**4 x 4 x 4**

2.4 2D Swin Transformer Training and Testing

2.4.1 2D Swin Transformer Network Structure

For 2D needle feature segmentation on the reformatted 2D image, we adapted the 2D Swin Transformer network for semantic segmentation. The network was implemented using Keras and PyTorch frameworks [52]. **Figure 2-6** depicts the architecture of the 2D Swin Transformer. Firstly, each input 2D image was split into non-overlapping patches by a patch-splitting module. These patches were then linearly embedded into a sequence of 1D tokens, which served as the input for the transformer layers. **Figure 2-6 (a)** shows the network architecture that comprised multiple stacked Swin Transformer blocks, which were responsible for capturing contextual information. **Figure 2-6 (b)** shows the structure of the Swin Transformer blocks which included (1) multi-head self-attention modules with regular and shifted windowing; (2) Multi-Layer Perceptron (MLP); (3)

Layer Normalization (LN); At the final stage of the architecture, a semantic segmentation head was used to generate the pixel-wise semantic labels.
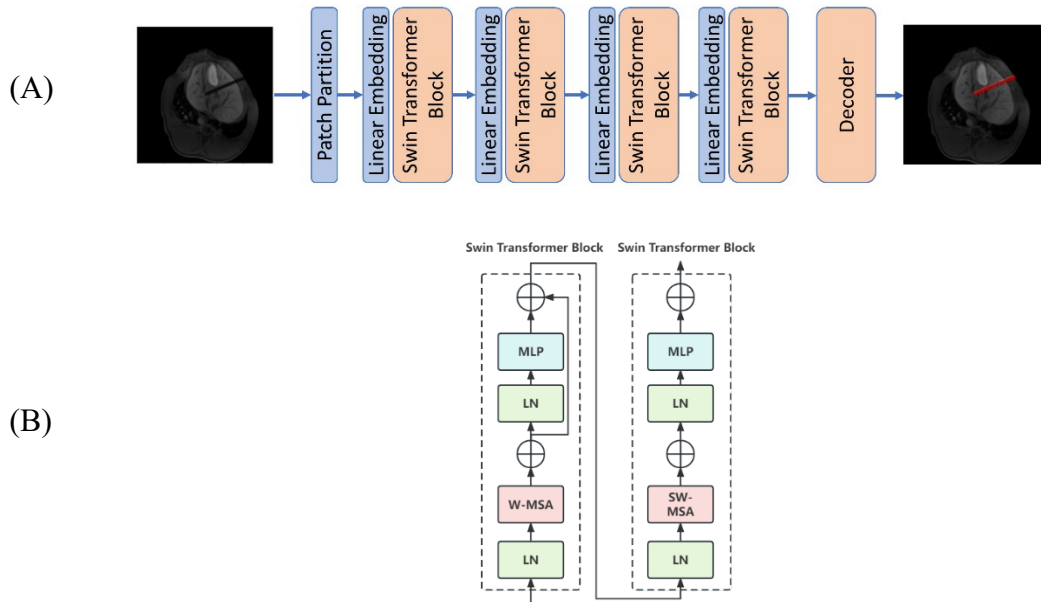
Figure 2-6 (a) The architecture of the 2D Swin Transformer. (b)Two successive Swin Transformer Blocks. W-MSA and SW-MSA are multi-head self-attention modules with regular and shifted windowing configurations, respectively.

2.4.2 2D Swin Transformer Training and Testing

The 2D Swin Transformer was first pre-trained with the 2D GRE images collected during the insertion stage of the experiment and then fine-tuned using the 2D reformatted images generated from the 49 3D T1w-VIBE images **(Table 2-2)**. To train the 2D Swin Transformer, we employed Weighted Dice Loss[51] as the training loss metric. Specific training hyperparameters are reported in **Table 2-4**.

To ensure the reliability of the 2D reformatted images used for cross-validation, the 3D needle feature segmentation generated by the 3D Swin Transformer served as the basis

for generating the 2D reformatted images. Post-processing modules in the pipeline were applied to the 3D needle feature segmentation to generate the 2D reformatted image plane that passes through the main axis of the predicted 3D needle segmentation.

To assess the effectiveness of the 2D Swin Transformer in needle feature segmentation, the study conducted seven-fold cross-validation. Each fold utilized 420 2D radial GRE images for pre-training and 42 2D reformatted images (from 3D T1w-VIBE) for fine-tuning. All images in the training dataset underwent 15-fold image augmentation. The Intersection over Union (IoU) score, ranging from 0 to 1, was used as a metric to measure the experimental results.

Table 2-4 2D Swin Transformer training parameters

	Pre-training (2D radial GRE dataset)	Fine-tuning (Reformatted 2D Dataset)
Batch size	8	8
Iteration	20000	10000
Weight decay	0.0001	0.0001
Learning rate	0.01	0.01
Window size	4 x 4	4 x 4

2.5 Evaluation of the Automatic 3D Needle Localization Pipeline

The needle localization pipeline takes intra-procedural 3D MR images as input and automatically locates the interventional needle feature tip coordinates and axis orientation in 3D. The pipeline was built in 3D Slicer[46] with SlicerIGT[53].

For pipeline performance evaluation, the needle tip and axis localization accuracy and image plane realignment accuracy were evaluated with the metrics shown in **Figure 2-7**. For 3D needle feature tip and axis localization evaluation, ε_{tip} (Euclidean distance between the predicted needle tip and reference needle tip) and α (The angle between the predicted needle axis and needle axis reference) were calculated. For 2D image plane realignment evaluation, δ_{tip} (Euclidean distance between the reformatted 2D image plane and the reference needle tip), δ_{entry} (Euclidean distance between the reformatted 2D image plane and the reference needle entry point), θ (The angle between the reference needle axis and the reformatted 2D image plane) was calculated.
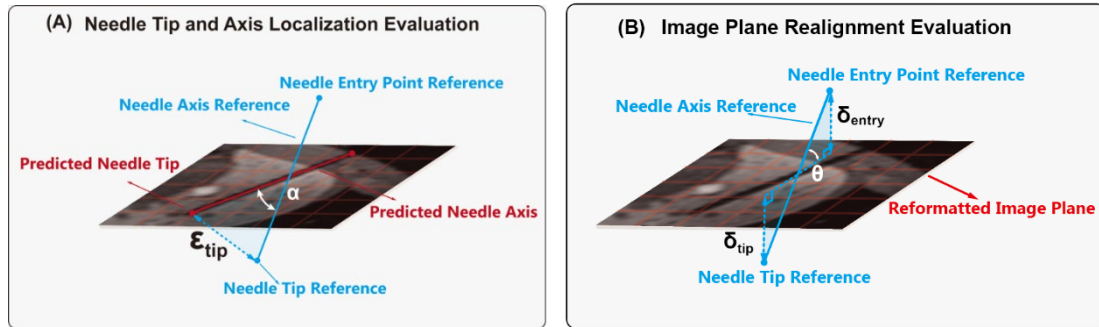
Figure 2-7 (A) Diagram of the needle tip and axis localization evaluation metrics. δ_{tip} **: Euclidean distance between the reformatted 2D image plane and the reference needle tip.** δ_{entry} **: Euclidean distance between the reformatted 2D image plane and the reference needle entry point.** θ **: The angle between the reference needle axis and the reformatted 2D image plane. (B) Diagram of image plane realignment evaluation metrics...** ε_{tip} **: Euclidean distance between the predicted needle tip and reference needle tip.** α **: The angle between the predicted needle axis and needle axis reference.**

Chapter 3 Results

3.1 3D Swin UNETR Evaluation

Figure 3-1 shows an example output of Swin UNETR. **Figure 3-1 (a)** depicts the input 3D T1w-VIBE MR images in axial, coronal, and sagittal views. **Figure 3-1(b)** depicts the 3D T1w-VIBE MR images overlaid with the predicted masks generated by the 3D Swin UNETR (red) and reference segmentation mask (blue). **Figure 3-1(c)** depicts the predicted 3D needle feature segmentation (red) overlaid with the reference 3D needle feature segmentation (blue) in the 3D view. The post-processing module in the pipeline calculates the volume of all the objects in the segmentation results and removes all the small objects as the needle feature segmentation typically has the largest volume. **Figure 3-1 (d)** depicts the predicted 3D needle feature segmentation (red) after post-processing overlaid with reference 3D needle feature segmentation (blue).
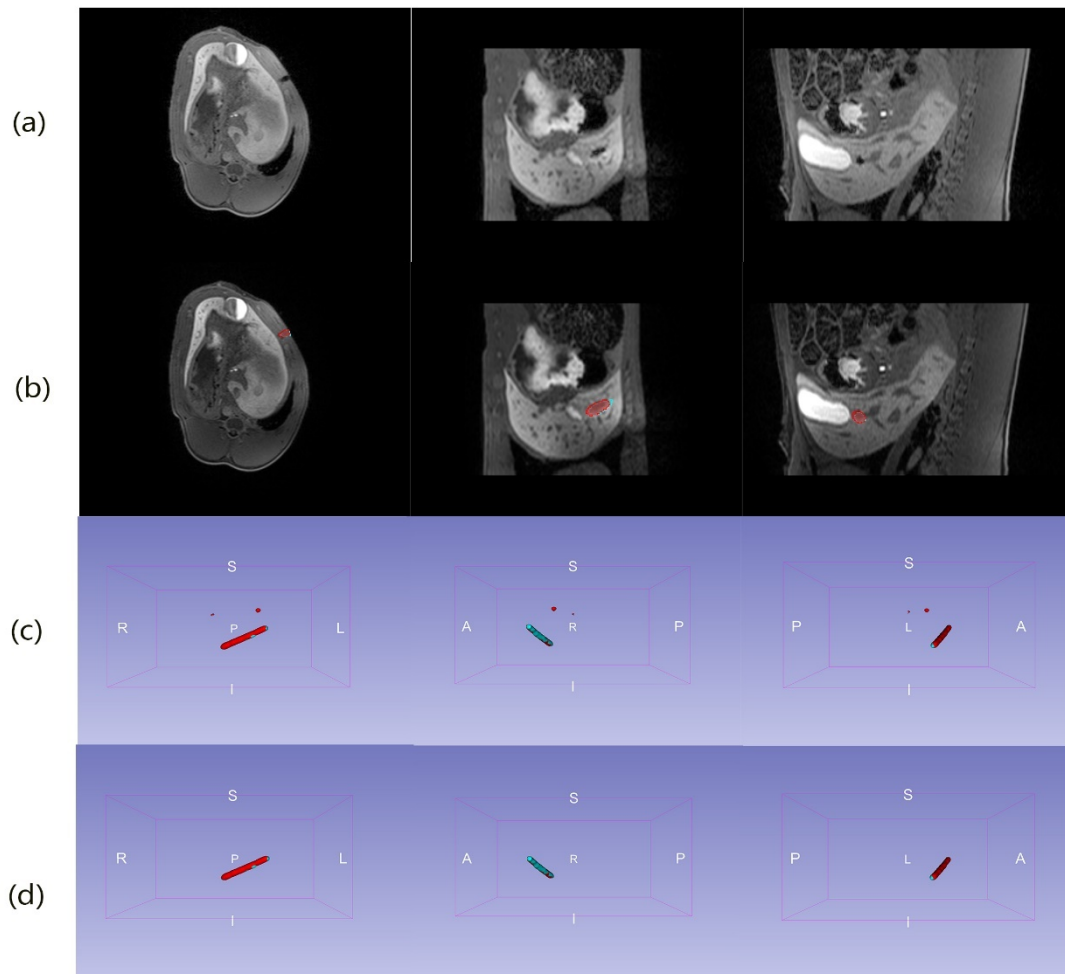
Figure 3-1 Segmentation results of the 3D Swin UNETR before and after post-processing (a) Input 3D T1-VIBE image. (b) Segmentation mask overlaid with input 3D T1-VIBE image. (c) Predicted 3D needle feature segmentation (red) overlaid with the reference 3D needle feature segmentation (blue) (d) Predicted 3D needle feature segmentation (red) after post-processing overlaid with the reference 3D needle feature segmentation (blue). For (c) and (d) the needle feature segmentations are displayed in 3D space using subject coordinates (R/L: right/left A/P: anterior/posterior, S/I: superior/inferior).

For the 49 instances in the seven-fold cross-validation, the average inference time of the 3D Swin UNETR was 2.14 sec per 3D volume on one NVIDIA RTX A6000 GPU card. **Table 3-1** summarizes the accuracy and precision of the 3D needle feature segmentation produced by 3D Swin UNETR, both with and without post-processing. The performance is quantified in terms of Dice Score, false positive rate, false negative

32

rate, true positive rate, and true negative rate. The mean and standard deviation (SD), median and interquartile range (IQR), and maximum values (i.e., worst-case performance) were reported.

In the 7-fold cross-validation involving the 49 3D T1w-VIBE images, the needle feature detection success rate was 100%. Of these 49 images, 34 contain segmentation results of multiple 3D objects. The post-processing modules succeeded in removing all false-positive segmentation objects while preserving the segmentation object that corresponded to the needle feature across all 49 instances.

Table 3-1 shows that after post-processing, the median Dice Score increased from 0.8042 to 0.8241. The median false-positive rate, after post-processing, decreased from 0.0026% to 0.0011%. This reduction occurred as a result of the false-positive removal module successfully removing all false-positive segmentation objects from the segmentation results. The false positive rate remains above zero after post-processing as the false-positive due to persisting false-positive segmentations within the largest object. Concurrently, the false negative rate increased as in some instances the needle feature segmentation comprised several smaller objects, and post-processing led to the removal of parts of the needle feature segmentation.

Table 3-1 3D needle feature segmentation accuracy of the 3D Swin UNETR. w/o pp: without post-processing. w/ pp: with post-processing.

	Mean	SD	Median	IQR	Max
Dice Score	w/o pp: 0.7768 w/ pp: 0.7771	w/o pp: 0.1010 w/ pp: 0.1156	w/o pp: 0.8042 w/ pp: 0.8241	w/o pp: 0.1050 w/ pp: 0.1479	w/o pp: 0.8985 w/ pp: 0.8989
False Positive Rate	w/o pp: 0.0038 w/ pp: 0.0023	w/o pp: 0.0051 w/ pp: 0.0030	w/o pp: 0.0026 w/ pp: 0.0011	w/o pp: 0.0039 w/ pp: 0.0032	w/o pp: 0.0325 w/ pp: 0.0128
False Negative Rate	w/o pp: 0.0114 w/ pp: 0.0121	w/o pp: 0.0080 w/ pp: 0.0084	w/o pp: 0.0104 w/ pp: 0.0106	w/o pp: 0.0117 w/ pp: 0.0132	w/o pp: 0.0333 w/ pp: 0.0334
True Positive Rate	w/o pp: 0.0290 w/ pp: 0.0284	w/o pp: 0.0200 w/ pp: 0.0202	w/o pp: 0.0237 w/ pp: 0.0235	w/o pp: 0.0179 w/ pp: 0.0173	w/o pp: 0.0886 w/ pp: 0.0885
True Negative Rate	w/o pp: 99.955 w/ pp: 99.957	w/o pp: 0.0272 w/ pp: 0.0260	w/o pp: 99.963 w/ pp: 99.965	w/o pp: 0.0240 w/ pp: 0.0238	w/o pp: 99.991 w/ pp: 99.991

3.2 2D Swin Transformer Evaluation

Figure 3-2 shows the segmentation results of the 2D Swin Transformer. **Figure 3-2(a)** depicts the input 2D reformatted MR images generated by the slice realignment module in the pipeline. **Figure 3-2(b)** depicts the 2D reformatted MR images overlaid with the predicted needle feature segmentation mask (red) from the 2D Swin Transformer. **Figure 3-2(c)** displays 2D reformatted MR images overlaid with the reference needle feature segmentation mask (blue).

(a) Original 2D reformatted image **(b) Predicted segmentation mask** **(c) Reference segmentation mask**
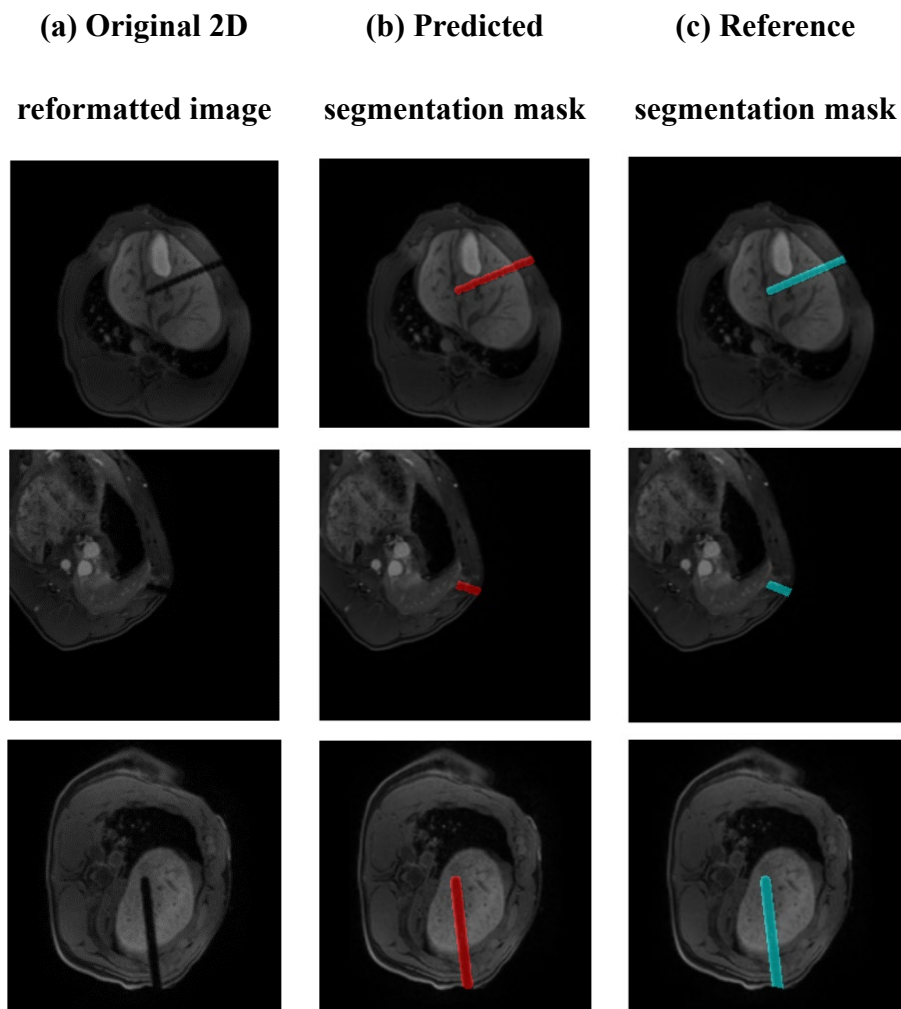
Figure 3-2 Example needle segmentation results on 2D reformatted MR images generated from the slice realignment module in the pipeline. (a) Original input 2D reformatted image generated by the pipeline. (b) Segmentation mask of the needle feature generated by the 3D Swin UNETR. (c) Reference needle feature segmentation mask.

For the 49 instances in the seven-fold cross-validation, the average inference time of the 2D Swin Transformer was 11 microseconds per image on one NVIDIA RTX A6000 GPU card. The testing result for the 2D needle feature segmentation of the 2D Swin Transformer model with and without fine-tuning are shown in **Table 3-2**. The non-fine-tuned model achieved a median Dice Score of 0.8940 and the fine-tuned model achieved a median Dice Score of 0.9261.

Table 3-2 2D Swin Transformer segmentation accuracy. w/o ft: without post-pro fine-tuning. w/ ft: with fine-tuning.

		Mean	SD	Median	IQR	Max
Dice Score	w/o ft:	0.8692	0.0709	**0.8940**	0.0786	0.9464
	w/ ft:	0.9092	0.0507	**0.9261**	0.0443	0.9673
False Positive Rate	w/o ft:	0.1248	0.0654	**0.1129**	0.0099	0.2853
	w/ ft:	0.0722	0.0429	**0.0686**	0.0732	0.1968
False Negative Rate	w/o ft:	0.0417	0.0402	**0.0290**	0.0320	0.2014
	w/ ft:	0.0462	0.0417	**0.0275**	0.0534	0.1632
True Positive Rate	w/o ft:	0.6537	0.3372	0.5874	0.3326	1.5600
	w/ ft:	0.6492	0.3214	0.5914	0.300%	1.4923
True Negative Rate	w/o ft:	99.179	0.3745	99.259	0.4394	99.789
	w/ ft:	99.232	0.3553	99.289	0.3799	99.763

3.3 Automatic Needle Localization Pipeline Evaluation

For needle feature localization, we performed 7-fold cross-validation on 49 3D T1w-VIBE images. In 49 T1w-VIBE images, the range of needle insertion depth was 1.94-12.26 cm. The range of needle insertion angle (angle between the needle and axial plane) was -87.64° to 2.23°. The accuracy of 3D needle localization was assessed by (A) needle tip and axis localization performance and (B) reformatted image plane localization performance.

(A)

(B)
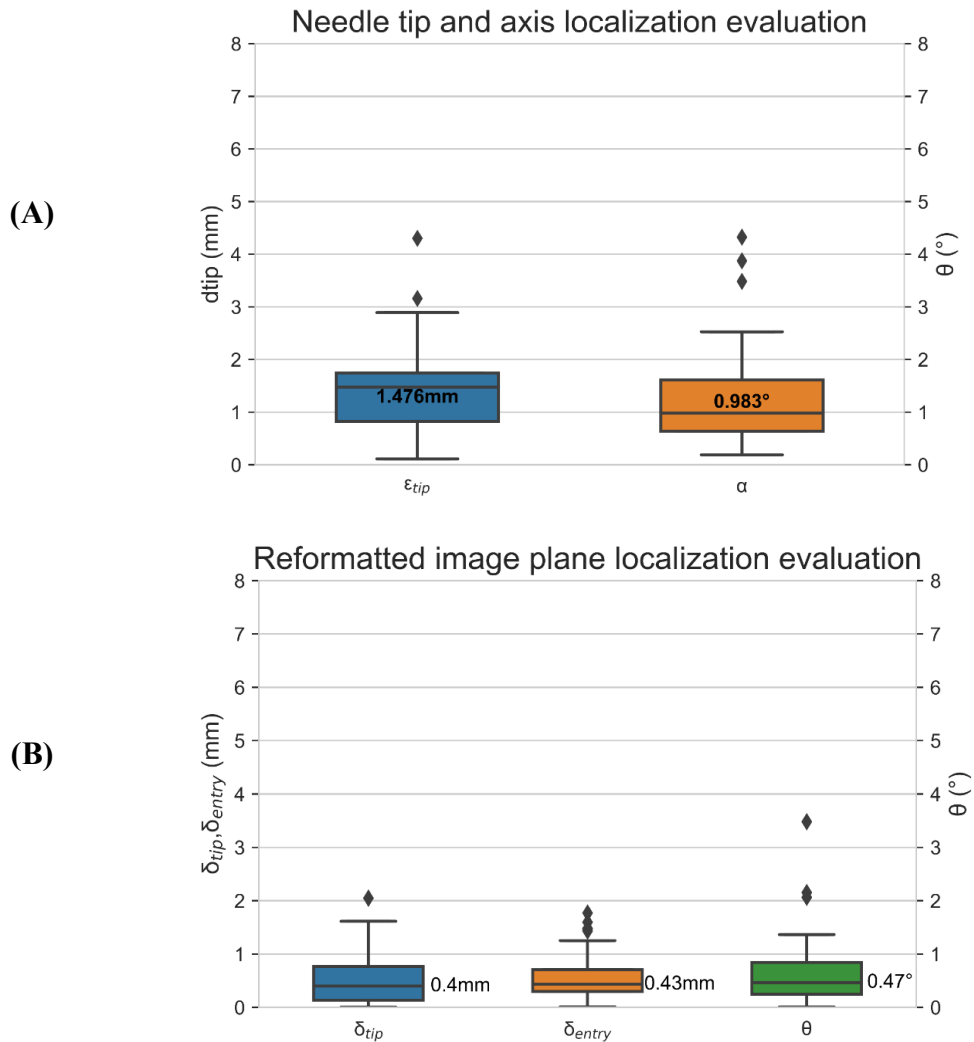
Figure 3-3 Evaluation of the pipeline with both 3D Swin UNETR and 2D Swin Transformer. (A)Box plots of needle feature tip and axis localization results. (B)Box plots of reformatted image plane localization evaluation results. The numbers reported in the plots are the medians of the results.

For the pipeline utilizing both 3D Swin UNETR and 2D Swin Transformer, the average computation time for localizing the needle feature tip and axis was approximately 6 seconds per instance from start to end. The needle detection was 100% successful in the 49 instances of the seven-fold cross-validation. **Figure 3-3** shows needle localization results of the pipeline from seven-fold cross-validation. For needle tip and axis localization, median ε_{tip}, which is the distance between the predicted needle tip and reference needle tip, was 1.476 mm (1.09 pixels). The median α, which is the angle between the predicted needle axis and the reference needle axis, was 0.983°. For reformatted image plane localization, median θ, which is the angle between the reference needle axis and the reformatted 2D image plane, was 0.47°. Representative 3D needle tip and axis localization results with oblique axial 2D reformatted image planes generated by the pipeline as reference are shown in **Figure 3-4.** These results showed accurate 3D needle localization using the automatic pipeline.

(a)3D Swin UNETR

segmentation results

and oriented bounding

box

(b) 2D Swin

Transformer

segmentation results

and 2D needle tip and

needle axis

(c) Needle tip and axis

localization results in 3D
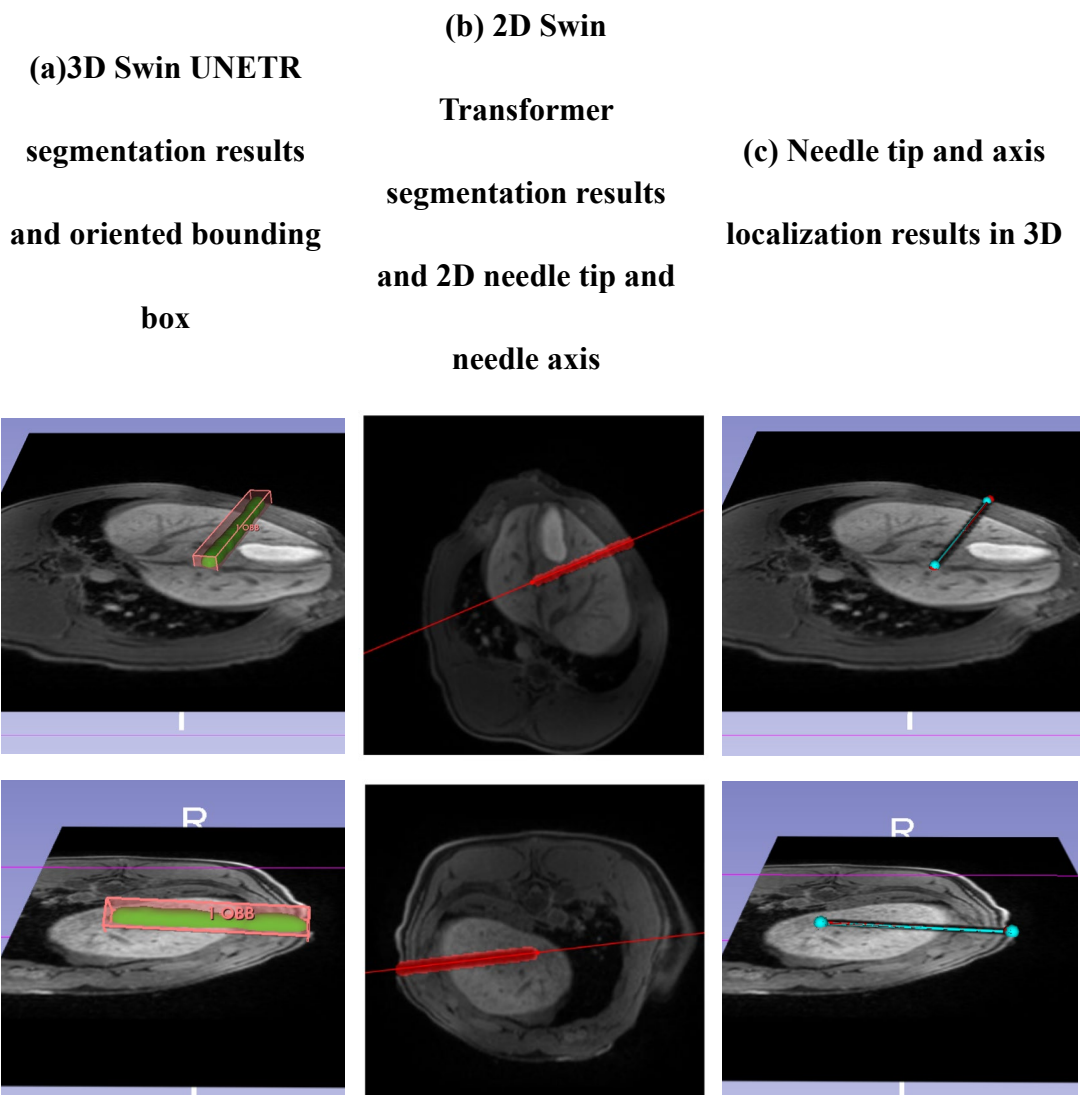
Figure 3-4 Example needle localization by the automatic pipeline. (a) 3D segmentation results of the 3D Swin UNETR and the oriented bounding box on the 2D reformatted image plane generated by the pipeline. (b) 2D Segmentation mask of the needle feature generated by the 2D Swin Transformer and needle axis on the 2D reformatted image plane. (c) Predicted needle tip and axis generated by the pipeline (red) and reference needle tip and axis (blue) in 3D. OBB: Oriented Bounding Box.

3.3.1 Comparison of 3D Needle Localization Pipelines with Different Structures

To support the design choices in our coarse-to-fine pipeline, we performed an ablation study where the proposed pipeline was compared with a pipeline that only employs the 3D Swin Transformer and predicted the needle location based on the 3D segmentation

results generated by the 3D Swin Transformer (without using a 2D Swin Transformer).

Figure 3-5 shows the structure of the needle localization pipeline that only utilizes 3D Swin UNETR. The pipeline takes 3D T1w-VIBE images as input and utilizes 3D UNETR for 3D needle feature segmentation. The segmentation output then undergoes small volume removal to eliminate false positives. Afterward, a bounding box is fitted to the 3D needle feature, with the main axis of the bounding box serving as the needle axis. The two intersections of the 3D needle feature segmentation and needle axis were designated as the needle tip and needle entry point. These two points are differentiated based on the average grayscale value surrounding them, with the one possessing higher grayscale values being considered the needle tip and the one with lower average grayscale values regarded as the needle entry point.
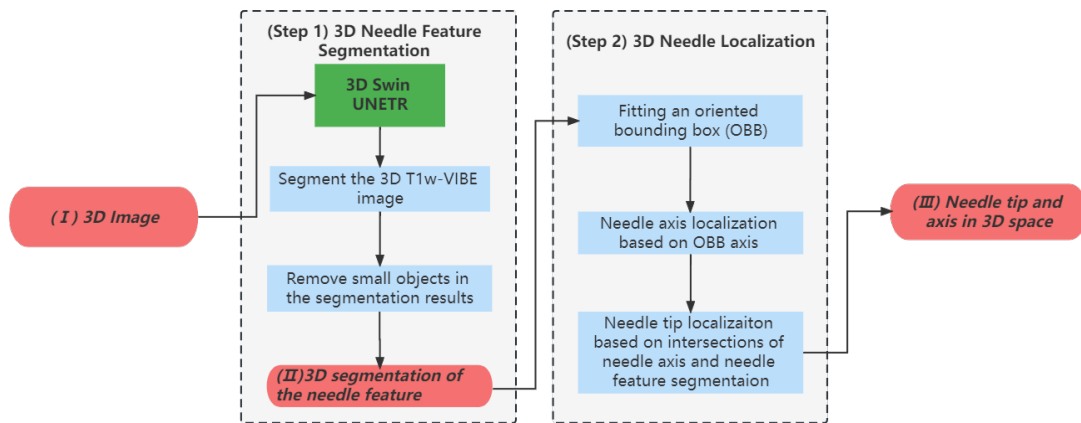
Figure 3-5 Structure of the pipeline that only employs 3D Swin Transformer.

Compared to the pipeline proposed in Chapter 2, the pipeline that only employs 3D Swin UNETR lacks correction from 2D fine segmentation generated by the 2D Swin Transformer network. Consequently, although the post-processing modules can remove

41

false-positives objects in segmentation results, false negatives (under-segmentation) persist, leading to more significant errors in the localization of the needle tip. However, both pipelines yielded the same image plane realignment outcomes because they both use the 3D needle feature segmentation for localizing the 2D reformatted image plane. On the other hand, compared with the pipeline proposed in Chapter 2 which has a computation time of 6 seconds, this pipeline achieves a computation time of about 4 seconds as it has fewer steps for needle localization.

Table 3-3 summarizes the needle tip and axis localization results of the proposed pipeline and the pipeline with only 3D Swin UNETR. By applying a 2D Swin Transformer in the pipeline, median ε_{tip} of decreased from 1.902 mm to 1.476 mm, and maximum ε_{tip} decreased from 37.083 mm to 4.430 mm. There is no notable improvement in needle axis localization after applying the 2D Swin Transformer network.

Table 3-3 3D needle localization accuracy. w/o 2D: Pipeline without 2D Swin Transformer network. w/ 2D: Pipeline with 2D Swin Transformer network.

	Mean		SD		Median		IQR		Max	
ε_{tip} (mm)	w/o 2D 4.426	w/ 2D 1.447	w/o 2D 7.727	w/ 2D 0.811	w/o 2D 1.902	w/ 2D 1.476	w/o 2D 1.430	w/ 2D 0.918	w/o 2D **37.083**	w/ 2D **4.430**
α (degree)	w/o 2D 1.073	w/ 2D 1.254	w/o 2D 0.712	w/ 2D 0.909	w/o 2D 0.921	w/ 2D 0.9825	w/o 2D 0.798	w/ 2D 0.976	w/o 2D 4.16	w/ 2D 4.324

Figure 3-6 shows the distribution of needle localization results of the pipeline with both 3D Swin UNETR and 2D Swin Transformer network compared with the pipeline utilizing only the 3D UNETR. The result of the Wilcoxon signed rank test indicates that the needle tip localization results were significantly improved by employing a 2D Swin Transformer in the pipeline. On the other hand, there is no significant difference in needle axis localization results generated by the two pipelines.
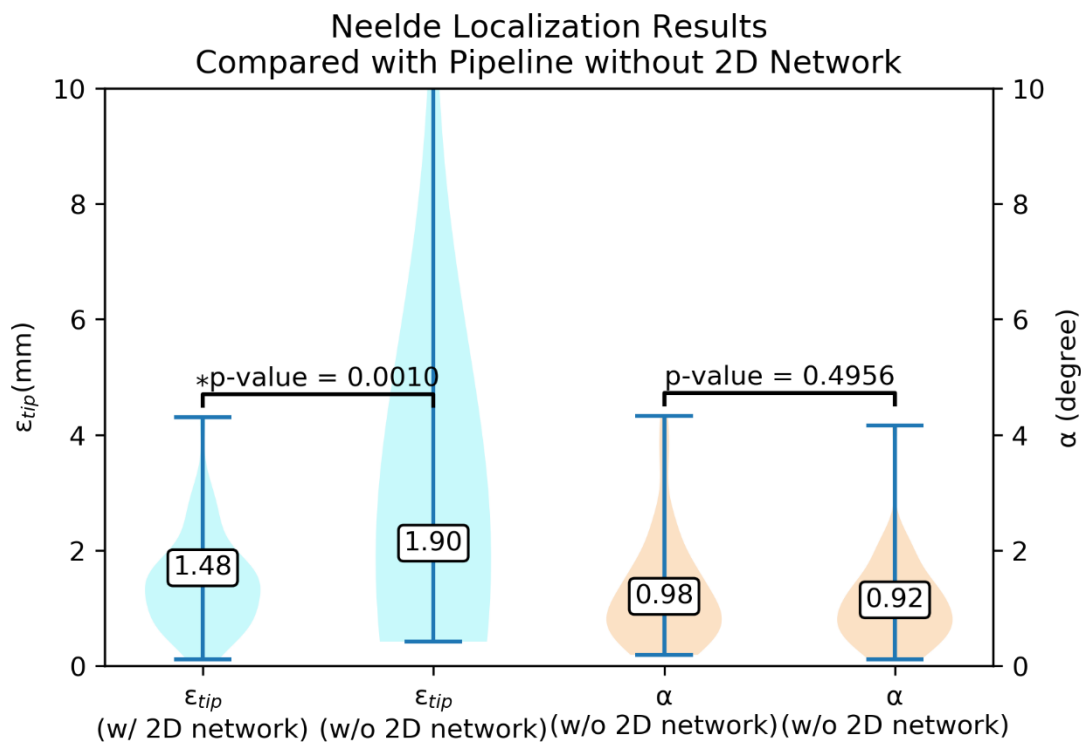
Figure 3-6 Violin plots of needle tip localization results(blue) and needle axis localization results(right) compared with a pipeline using only 3D UNETR.
In the pair-wise comparison, the Wilcoxon signed rank test shows p-value=0.0010 for needle tip localization results and p-value=0.4956 for needle axis localization results. * indicates Wilcoxon signed rank test with p<0.01. w/ 2D network: with 2D network. w/o 2D network: without 2D network.

Figure 3-7 Scatter plot of ε_{tip} and α for insertion depth in the proposed pipeline and in the pipeline that only employs the 3D Swin UNETR. **Figure 3-7 (A)** shows that 2D

Swin UNETR increases the pipeline's performance in needle tip localization by removing the outliers in needle tip localization, as the under-segmentation of the 3D Swin Transformer can be compensated by applying the 2D Swin Transformer with higher precision. **Figure 3-7 (B)** shows that the needle axis localization results of the two pipelines are comparable. Furthermore, the needle axis localization error decreases with the needle insertion depth as when needle insertion depths are small, the inaccuracy of the 3D needle segmentation predictions tends to have a greater effect on the needle axis localization results.
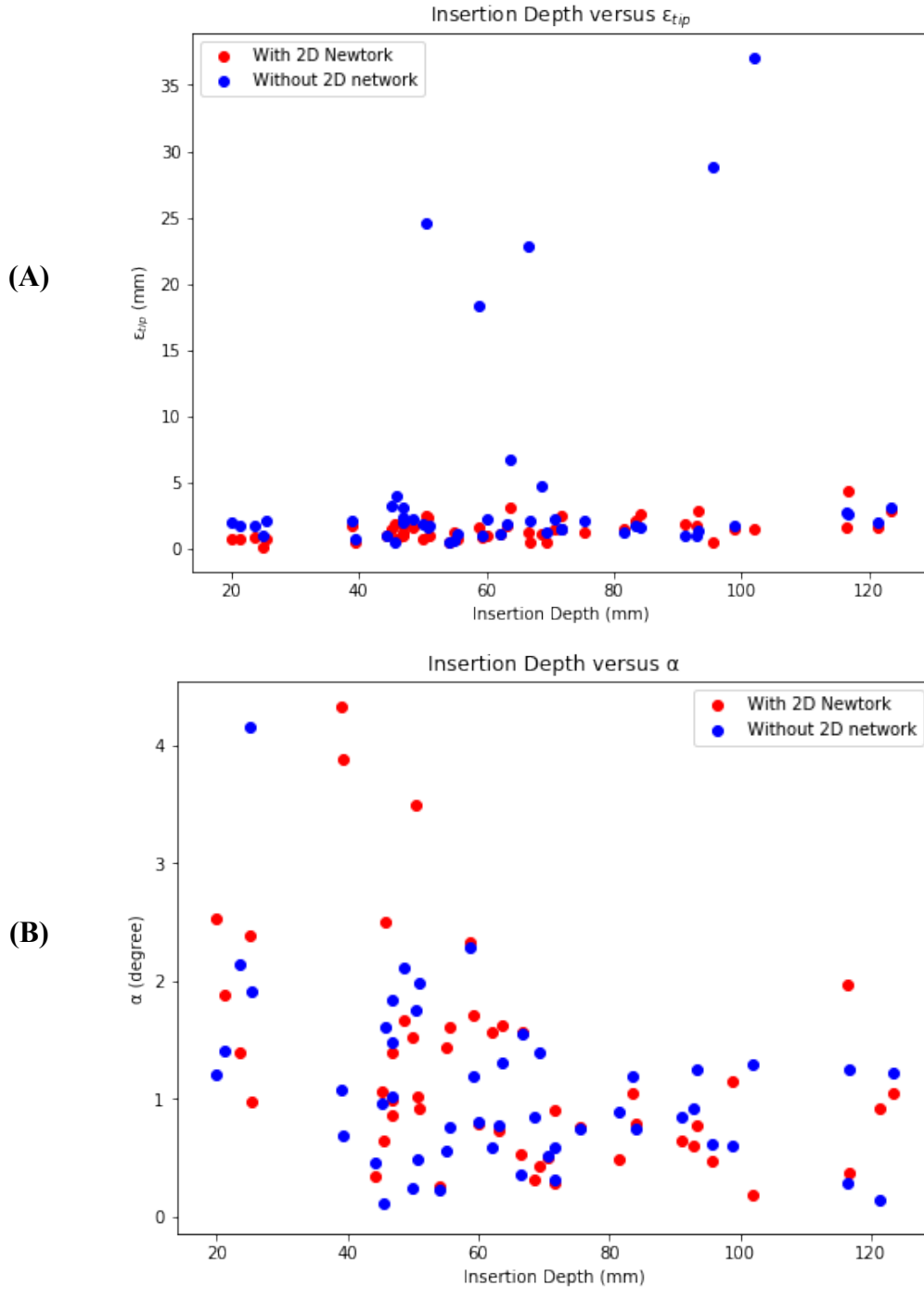
Figure 3-7 Scatter plot of (A) ε_{tip} and (B) α with respect to insertion depth in the proposed pipeline and in the pipeline that only employs the 3D Swin UNETR.

3.3.2 Comparison of 3D Needle Localization Pipelines Results with Human Intra-Reader Variation

During the needle tip and axis annotation process, the needle tip and entry point were

marked on the 3D Slicer, and the process was repeated twice with a wash-out period of two weeks in between sessions. The needle localization difference between the two annotations was recorded as a human intra-reader variation. ε_{tip} which is the needle tip distance between the predicted needle tip and reference needle tip, and α which is the angle between the predicted needle axis and reference needle axis were compared with the human intra-reader variation. The distribution of the needle localization results from seven-fold cross-validation and human intra-reader variation of all the images are summarized and compared using violin plots in **Figure 3-6**. Based on Wilcoxon signed rank test, there is a significant difference in ε_{tip}, and no significant difference in $\boldsymbol{\alpha}$.
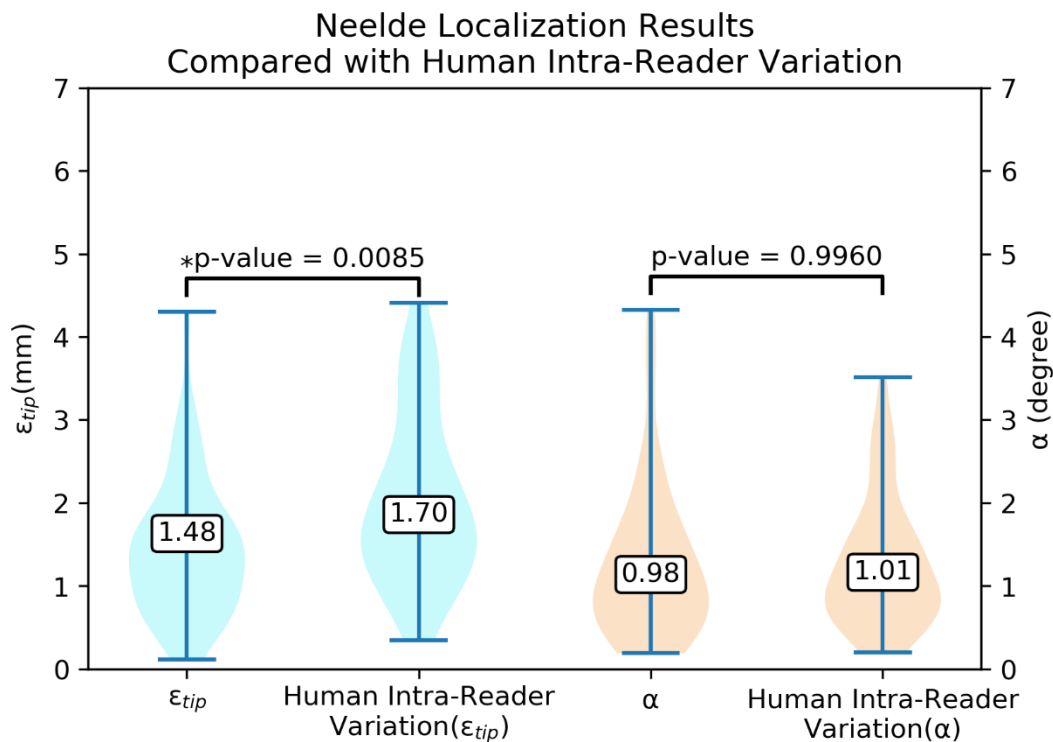

Figure 3-8 Violin plots of needle tip localization results(blue) and needle axis localization results(right) compared with the human intra-reader variations.
In the pair-wise comparison, the Wilcoxon signed rank test shows p-value=0.0085 for needle tip localization results and p-value=0.9960 for needle axis localization results. * indicates Wilcoxon signed rank test with p<0.01.

Chapter 4 Discussion

In this study, we have developed an automatic 3D needle localization pipeline for MRI-guided interventions. The pipeline incorporated a 3D Swin UNETR and a 2D Swin Transformer. First, the 3D Swin UNETR produced an initial coarse segmentation of the needle feature on intra-procedural 3D MR images. Next, the post-processing modules generated reformatted 2D images that aligned with the 3D needle axis. The reformatted 2D images were then sent to the 2D Swin Transformer for fine segmentation. The fine 2D segmentation of the needle feature was then used to calculate the 2D coordinates of the needle tip, entry point, and needle axis. The 2D coordinates were converted back to 3D space and produced 3D needle localization results.

We performed seven-fold cross-validation with a total of 49 *in vivo* 3D interventional MR images. The cross-validation results demonstrated consistent 3D needle localization accuracy, with a median needle tip localization error of 1.476 mm (1.09 pixels) and a median needle axis localization error of 0.98°, which was comparable to human reader variability. This level of accuracy is sufficient for liver biopsy since clinically-relevant lesions typically have a diameter of 5-10 mm[54,55]. The pipeline also achieved real-time 3D needle localization, with a processing time of around 6 seconds from start to finish.

As part of the pipeline, we adopted 3D Swin UNETR for 3D needle feature segmentation. The 3D Swin UNETR achieved a median Dice Score of 0.8042 in 3D

needle feature segmentation. The 3D UNETR trained on a relatively small 3D MR image dataset (~40 volumetric images) and achieved the accuracy mentioned above with its transformer-based network design to capture the long-range information in the 3D image and its adoption of the pre-trained model weights generated by the self-supervised pre-training on the CT images[42]. The reformatted 2D image plane realignment results generated based on the 3D segmentation result of the 3D Swin UNETR achieved a median θ (angle between the predicted reformatted image plane and reference needle axis) of 0.47° which is adequate for guiding image plane realignment during interventional procedures.

We also developed a 2D Swin Transformer for 2D needle feature segmentation on the 2D reformatted image plane generated by the pipeline. The 2D Swin Transformer achieved a median Dice Score of 0.8940 after pre-training with the 2D single-slice GRE images of the 2D needle feature collected during the interventional experiments and a median Dice Score of 0.9261 after fine-tuning with the reformatted 2D images generated from 3D T1w-VIBE by the pipeline. By combining the 2D Swin Transformer and the 3D Swin UNETR, the pipeline's performance of needle tip localization improved from a median ε_{tip} (distance between the predicted needle tip to the reference needle tip) of 1.902 mm to 1.476 mm. During the interventional procedures, the 2D fine segmentation generated by the 2D Swin Transformer in the pipeline will be of great clinical significance in the final insertion stage and confirmation stage where the precise needle tip and axis localization are required.

49

There are limitations to this study. Firstly, due to the limitation of the size of the dataset, all the results shown here are from the cross-validation experiments. In the future, independent testing can be done to further assess the pipeline's performance. Secondly, in this study, the position of the needle tip and axis was estimated by the location of the needle feature tip and needle feature axis. However, a major limitation is that the discrepancy between the needle feature and the underlying physical needle position has not been addressed. The susceptibility difference between a physical needle and water causes magnetic field perturbation and MR signal dephasing. The signal void features caused by the susceptibility effect can have an irregular shape at the tip and the axis can be shifted from the physical needle axis. Previous studies have shown that the needle tip artifact can extend beyond the actual needle tip location and exhibit strong curvature in the direction of the B_0 magnetic field[56,57]. Nevertheless, the observed titanium needle feature void artifact displacements are typically 2 mm or less from the actual needle location, smaller than the radius of the clinically significant tumor size of approximately 4.9 mm[56]. Thirdly, the pipeline design can be further improved. The post-processing module employed in the pipeline removes the false-positive objects by calculating the volume of each object and removing the small one. In the future, the false-positive removal modules could incorporate relative position information of the needle and surrounding tissue to enhance performance.

Chapter 5 Future Work and Conclusion

In summary, this study achieved the first aim of developing neural networks for 3D and 2D needle feature segmentation on intra-procedural MRI by the implementation of 3D Swin UNETR and 2D Swin Transformer. Based on the results from cross-validation, the 3D Swin achieved a median Dice Score of 0.8042 in 3D needle feature segmentation on 3D T1w-VIBE images collected during the interventional experiments. The 2D Swin Transformer achieved a median Dice Score of 0.9261 in 2D needle feature segmentation on the 2D reformatted images generated by the pipeline. The second aim of developing an automatic deep learning-based pipeline for 3D needle localization on intraprocedural MRI was achieved by the implementation of the proposed pipeline. The cross-validation results show that the proposed pipeline achieved real-time and accurate 3D needle localization with a computation time of 6 seconds, median needle tip localization error of 1.476 mm, and median needle axis localization error of 0.98 degrees which is comparable to human intra-reader variation.

5.1 Future Work

Firstly, the time efficiency of the pipeline will be improved in the future. The average computation time for 3D needle localization in each instance was about 6 sec. In the pipeline, the inference time for 3D Swin UNETR was about 2 seconds and the inference time for 2D Swin Transformer was about 11 microseconds. Therefore, there is still room for improvement in accelerating the processing time of data loading and transferring by optimizing the pipeline design. The optimal processing time for needle localization in

each instance will be around 2-3 seconds, which is within the time constraints of interventional procedures. Secondly, to address the challenges of limitation of the size of the 3D dataset, needle feature synthesis in 3D MR images will be performed to expand the dataset of the 3D images[47]. The synthesized images can be utilized for training and testing the 3D neural network and the pipeline. Thirdly, MRI scan plane control modules will be implemented by utilizing the realignment image plane localization results generated by the pipeline. The scan plane control module in the scanner will enable real-time scan plane updates during the interventional experiments to streamline the workflow. Fourthly, the discrepancy between the physical needle tip and needle feature tip will be taken into account by applying physical needle localization methods preciously studied[47]. Finally, additional *in vivo* interventional experiments will be conducted to collect more data for training and also to further validate the pipeline's performance.

5.2 Conclusion

The work in this thesis developed a pipeline for real-time automatic 3D needle localization in MRI-guided interventions. The pipeline has a coarse-to-fine structure where it adopts 3D Swin UNETR for coarse segmentation of the 3D needle feature and 2D Swin Transformer for fine segmentation of the needle feature in the 2D reformatted image plane. The proposed pipeline achieved real-time and accurate 3D needle localization and thus further expanded clinical applications of MRI-guided percutaneous interventions.

References

1. Cleary K, Peters TM. Image-Guided Interventions: Technology Review and Clinical Applications. *Annual Review of Biomedical Engineering*. 2010;12(1):119-142. doi:10.1146/annurev-bioeng-070909-105249

2. Maybody M, Stevenson Carsten, Solomon SB. Overview of Navigation Systems in Image-Guided Interventions. *Techniques in Vascular and Interventional Radiology*. 2013;16(3):136-143. doi:10.1053/j.tvir.2013.02.008

3. Wein W, Brunke S, Khamene A, Callstrom MR, Navab N. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis*. 2008;12(5):577-585. doi:10.1016/j.media.2008.06.006

4. Needle Visualization in Ultrasound-Guided Regional Anesthesia: Challenges and Solutions | Regional Anesthesia & Pain Medicine. Accessed April 19, 2023. https://rapm.bmj.com/content/33/6/532.abstract

5. Douglas BR, Charboneau JW, Reading CC. ULTRASOUND-GUIDED INTERVENTION: Expanding Horizons. *Radiologic Clinics of North America*. 2001;39(3):415-428. doi:10.1016/S0033-8389(05)70289-X

6. Nodular hepatocellular carcinomas: detection with arterial-, portal-, and delayed-phase images at spiral CT. | Radiology. Accessed April 19, 2023. https://pubs.rsna.org/doi/abs/10.1148/radiology.202.2.9015062

7. Assessing tumor response after loco-regional liver cancer therapies: the role of 3D MRI. Accessed April 19, 2023. https://www.tandfonline.com/doi/epdf/10.1586/14737140.2015.978861?needAccess=true&role=button

8. Beyersdorff D, Winkel A, Hamm B, Lenk S, Loening SA, Taupitz M. MR imaging-guided prostate biopsy with a closed MR unit at 1.5 T: initial results. *Radiology*. 2005;234(2):576-581. doi:10.1148/radiol.2342031887

9. Campbell-Washburn AE, Tavallaei MA, Pop M, et al. Real-time MRI guidance of cardiac interventions. *Journal of Magnetic Resonance Imaging*. 2017;46(4):935-950. doi:10.1002/jmri.25749

10. Lazebnik RS, Lancaster TL, Breen MS, Lewin JS, Wilson DL. Volume registration using needle paths and point landmarks for evaluation of interventional MRI treatments. *IEEE Transactions on Medical Imaging*. 2003;22(5):653-660. doi:10.1109/TMI.2003.812246

11. Mehrtash A, Ghafoorian M, Pernelle G, et al. Automatic Needle Segmentation and Localization in MRI With 3-D Convolutional Neural Networks: Application to MRI-

Targeted Prostate Biopsy. *IEEE Transactions on Medical Imaging*. 2019;38(4):1026-1036. doi:10.1109/TMI.2018.2876796

12. Zhang J, Liu M, Shen D. Detecting Anatomical Landmarks From Limited Medical Imaging Data Using Two-Stage Task-Oriented Deep Neural Networks. *IEEE Transactions on Image Processing*. 2017;26(10):4753-4764. doi:10.1109/TIP.2017.2721106

13. Hushek SG, Martin AJ, Steckner M, Bosak E, Debbins J, Kucharzyk W. MR systems for MRI-guided interventions. *Journal of Magnetic Resonance Imaging*. 2008;27(2):253-266. doi:10.1002/jmri.21269

14. Kaye EA, Granlund KL, Morris EA, Maybody M, Solomon SB. Closed-Bore Interventional MRI: Percutaneous Biopsies and Ablations. *American Journal of Roentgenology*. 2015;205(4):W400-W410. doi:10.2214/AJR.15.14732

15. Arnolli MM, Hanumara NC, Franken M, Brouwer DM, Broeders IAMJ. An overview of systems for CT- and MRI-guided percutaneous needle placement in the thorax and abdomen. *The International Journal of Medical Robotics and Computer-Assisted Surgery*. 2015;11(4):458-475. doi:10.1002/rcs.1630

16. Morrison PR, Silverman SG, Tuncali K, Tatli S. MRI-guided cryotherapy. *Journal of Magnetic Resonance Imaging*. 2008;27(2):410-420. doi:10.1002/jmri.21260

17. Interventional Radiology Robot for CT and MRI Guided Percutaneous Interventions | SpringerLink. Accessed May 3, 2023. https://link.springer.com/chapter/10.1007/978-3-642-23623-5_18

18. Henken KR, Seevinck PR, Dankelman J, van den Dobbelsteen JJ. Manually controlled steerable needle for MRI-guided percutaneous interventions. *Med Biol Eng Comput*. 2017;55(2):235-244. doi:10.1007/s11517-016-1490-0

19. Kaiser WA, Pfleiderer SOR, Baltzer PAT. MRI-guided interventions of the breast. *Journal of Magnetic Resonance Imaging*. 2008;27(2):347-355. doi:10.1002/jmri.21276

20. Gignac PM, Kley NJ, Clarke JA, et al. Diffusible iodine-based contrast-enhanced computed tomography (diceCT): an emerging tool for rapid, high-resolution, 3-D imaging of metazoan soft tissues. *Journal of Anatomy*. 2016;228(6):889-909. doi:10.1111/joa.12449

21. Lee JW, Choi MH, Lee YJ, et al. Radiofrequency ablation for liver metastases in patients with gastric cancer as an alternative to hepatic resection. *BMC Cancer*. 2017;17(1):185. doi:10.1186/s12885-017-3156-1

22. Lindner NJ, Ozaki T, Roedl R, Gosheger G, Winkelmann W. Percutaneous radiofrequency ablation in osteoid osteoma. *The Journal of Bone & Joint Surgery*

British Volume. 2001;83-B(3):391-396. doi:10.1302/0301-620x.83b3.0830391

23. Nian-Long L, Bo Y, Tian-Ming C, et al. The application of magnetic resonance imaging-guided microwave ablation for lung cancer. *Journal of Cancer Research and Therapeutics*. 2020;16(5):1014. doi:10.4103/jcrt.JCRT_354_20

24. Clasen S, Pereira PL. Magnetic resonance guidance for radiofrequency ablation of liver tumors. *Journal of Magnetic Resonance Imaging*. 2008;27(2):421-433. doi:10.1002/jmri.21264

25. Hoffmann R, Rempp H, Schraml C, et al. Diffusion-weighted imaging during MR-guided radiofrequency ablation of hepatic malignancies: analysis of immediate pre- and post-ablative diffusion characteristics. *Acta Radiol*. 2015;56(8):908-916. doi:10.1177/0284185114545148

26. Izzo F, Granata V, Grassi R, et al. Radiofrequency Ablation and Microwave Ablation in Liver Tumors: An Update. *The Oncologist*. 2019;24(10):e990-e1005. doi:10.1634/theoncologist.2018-0337

27. Vogl TJ, Straub R, Eichler K, Woitaschek D, Mack MG. [Modern alternatives to resection of metastases--MR-guided laser-induced thermotherapy (LITT) and other local ablative techniques. *Ther Umsch*. 2001;58(12):718-725. doi:10.1024/0040-5930.58.12.718

28. Mohammadi AM, Schroeder JL. Laser interstitial thermal therapy in treatment of brain tumors – the NeuroBlate System. *Expert Review of Medical Devices*. 2014;11(2):109-119. doi:10.1586/17434440.2014.882225

29. Busse H, Kahn T, Moche M. Techniques for Interventional MRI Guidance in Closed-Bore Systems. *Topics in Magnetic Resonance Imaging*. 2018;27(1):9. doi:10.1097/RMR.0000000000000150

30. Mikaiel S, Simonelli J, Li X, et al. MRI-guided targeted needle placement during motion using hydrostatic actuators. *The International Journal of Medical Robotics and Computer-Assisted Surgery*. 2020;16(2):e2041. doi:10.1002/rcs.2041

31. Su H, Shang W, Cole G, et al. Piezoelectrically Actuated Robotic System for MRI-Guided Prostate Percutaneous Therapy. *IEEE/ASME Transactions on Mechatronics*. 2015;20(4):1920-1932. doi:10.1109/TMECH.2014.2359413

32. Wu D, Li G, Patel N, et al. Remotely Actuated Needle Driving Device for MRI-Guided Percutaneous Interventions. In: *2019 International Symposium on Medical Robotics (ISMR)*. ; 2019:1-7. doi:10.1109/ISMR.2019.8710176

33. Ladd ME, Erhart P, Debatin JF, Romanowski BJ, Boesiger P, McKinnon GC. Biopsy needle susceptibility artifacts. *Magnetic Resonance in Medicine*. 1996;36(4):646-651. doi:10.1002/mrm.1910360423

34. Schenck JF. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Medical Physics*. 1996;23(6):815-850. doi:10.1118/1.597854

35. Lewin JS, Petersilge CA, Hatem SF, et al. Interactive MR imaging-guided biopsy and aspiration with a modified clinical C-arm system. *American Journal of Roentgenology*. 1998;170(6):1593-1601. doi:10.2214/ajr.170.6.9609180

36. Interventional MR Imaging: Concepts, Systems, and Applications in Neuroradiology | American Journal of Neuroradiology. Accessed May 3, 2023. https://www.ajnr.org/content/20/5/735.short

37. Frahm C, Gehl HB, Melchert UH, Weiss HD. Visualization of magnetic resonance-compatible needles at 1.5 and 0.2 Tesla. *Cardiovasc Intervent Radiol*. 1996;19(5):335-340. doi:10.1007/BF02570186

38. Li X, Young AS, Raman SS, et al. Automatic needle tracking using Mask R-CNN for MRI-guided percutaneous interventions. *Int J CARS*. 2020;15(10):1673-1684. doi:10.1007/s11548-020-02226-8

39. Vaswani A, Shazier N, Parmar N, et al. Attention is All you Need. In: *Advances in Neural Information Processing Systems*. Vol 30. Curran Associates, Inc.; 2017. Accessed April 19, 2023. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

40. Liu Z, Lin Y, Cao Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. In; 2021:10012-10022. Accessed April 19, 2023. https://openaccess.thecvf.com/content/ICCV2021/html/Liu_Swin_Transformer_Hierarchical_Vision_Transformer_Using_Shifted_Windows_ICCV_2021_paper.html

41. Hatamizadeh A, Nath V, Tang Y, Yang D, Roth HR, Xu D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. In: Crimi A, Bakas S, eds. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Lecture Notes in Computer Science. Springer International Publishing; 2022:272-284. doi:10.1007/978-3-031-08999-2_22

42. Tang Y, Yang D, Li W, et al. Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis. In; 2022:20730-20740. Accessed April 19, 2023. https://openaccess.thecvf.com/content/CVPR2022/html/Tang_Self-Supervised_Pre-Training_of_Swin_Transformers_for_3D_Medical_Image_Analysis_CVPR_2022_paper.html

43. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, eds. *Medical*

Image Computing and Computer-Assisted Intervention – MICCAI 2015. Lecture Notes in Computer Science. Springer International Publishing; 2015:234-241. doi:10.1007/978-3-319-24574-4_28

44. Maurya A, Patil KD, Padhy R, Ramakrishna K, Krishnamurthi G. PARSE challenge 2022: Pulmonary Arteries Segmentation using Swin U-Net Transformer(Swin UNETR) and U-Net. Published online August 20, 2022. doi:10.48550/arXiv.2208.09636

45. Li X, Lee YH, Mikaiel S, Simonelli J, Tsao TC, Wu HH. Respiratory Motion Prediction Using Fusion-Based Multi-Rate Kalman Filtering and Real-Time Golden-Angle Radial MRI. *IEEE Transactions on Biomedical Engineering*. 2020;67(6):1727-1738. doi:10.1109/TBME.2019.2944803

46. Pieper S, Halle M, Kikinis R. 3D Slicer. In: *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*. ; 2004:632-635 Vol. 1. doi:10.1109/ISBI.2004.1398617

47. Li X, Lee YH, Lu DS, Tsao TC, Wu HH. Physics-Driven Mask R-CNN for Physical Needle Localization in MRI-Guided Percutaneous Interventions. *IEEE Access*. 2021;9:161055-161068. doi:10.1109/ACCESS.2021.3128163

48. Poch J, Villaescusa I. Orthogonal Distance Regression: A Good Alternative to Least Squares for Modeling Sorption Data. *J Chem Eng Data*. 2012;57(2):490-499. doi:10.1021/je201070u

49. Cao H, Wang Y, Chen J, et al. Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. In: Karlinsky L, Michaeli T, Nishino K, eds. *Computer Vision – ECCV 2022 Workshops*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2023:205-218. doi:10.1007/978-3-031-25066-8_9

50. Ulyanov D, Vedaldi A, Lempitsky V. Instance Normalization: The Missing Ingredient for Fast Stylization. Published online November 6, 2017. doi:10.48550/arXiv.1607.08022

51. Zhao R, Qian B, Zhang X, et al. Rethinking Dice Loss for Medical Image Segmentation. In: *2020 IEEE International Conference on Data Mining (ICDM)*. ; 2020:851-860. doi:10.1109/ICDM50108.2020.00094

52. Swin Transformer for Semantic Segmentation. Published online April 18, 2023. Accessed April 19, 2023. https://github.com/SwinTransformer/Swin-Transformer-Semantic-Segmentation

53. Ungi T, Lasso A, Fichtinger G. Open-source platforms for navigated image-guided interventions. *Med Image Anal*. 2016;33:181-186. doi:10.1016/j.media.2016.06.011

54. Moche M, Heinig S, Garnov N, et al. Navigated MRI-guided liver biopsies in a

closed-bore scanner: experience in 52 patients. *Eur Radiol*. 2016;26(8):2462-2470. doi:10.1007/s00330-015-4097-1

55. Parekh PJ, Majithia R, Diehl DL, Baron TH. Endoscopic ultrasound-guided liver biopsy. *Endosc Ultrasound*. 2015;4(2):85-91. doi:10.4103/2303-9027.156711

56. Song SE, Cho NB, Iordachita II, Guion P, Fichtinger G, Whitcomb LL. A Study of Needle Image Artifact Localization in Confirmation Imaging of MRI-guided Robotic Prostate Biopsy. *IEEE Int Conf Robot Autom*. 2011;2011:4834-4839. doi:10.1109/ICRA.2011.5980309

57. DiMaio SP, Kacher DF, Ellis RE, et al. Needle artifact localization in 3T MR images. *Stud Health Technol Inform*. 2006;119:120-125.