Physics-Driven Mask R-CNN for Physical Needle Localization in MRI-Guided Percutaneous Interventions

XINZHOU LI[1,2]**, YU-HSIU LEE**[3,4]**, DAVID S. LU**[1]**, TSU-CHIN TSAO**[3]**, (Senior Member, IEEE), AND HOLDEN H. WU**[1,2]**, (Member, IEEE)**

[1]Department of Radiological Sciences, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA
[2]Department of Bioengineering, University of California Los Angeles, Los Angeles, CA 90095, USA
[3]Department of Mechanical and Aerospace Engineering, University of California Los Angeles, Los Angeles, CA 90095, USA
[4]Department of Mechanical Engineering, National Taiwan University, Taipei 10617, Taiwan

Corresponding author: Holden H. Wu (holdenwu@mednet.ucla.edu)

ABSTRACT Discrepancies between the needle feature position on magnetic resonance imaging (MRI) and the underlying physical needle position could increase localization errors during needle-based targeting procedures in MRI-guided percutaneous interventions. This work aimed to develop a deep learning-based framework to automatically localize the physical needle position using only the needle features on MR images. Physics-based simulations were performed to generate single-slice and 3-slice images with needle features from a range of underlying needle positions and MRI parameters to form datasets for training single-slice and 3-slice Mask Region-Based Convolutional Neural Network (R-CNN) models for physical needle localization. *Ex vivo* tissue images were combined with simulated needle features for fine-tuning. Next, the physics-driven Mask R-CNN models were combined with a previously developed Mask R-CNN model for needle feature localization to form an automated framework to localize the physical needle. To test the accuracy of the proposed framework, both single-slice and 3-slice MRI data were acquired from needle insertion experiments in *ex vivo* tissue phantoms. Using the single-slice model, the proposed framework achieved sub-millimeter physical needle localization accuracy on single-slice images aligned with the needle. The fine-tuning step reduced in-plane physical needle tip localization error (mean±standard deviation) to 0.96±0.69 mm in *ex vivo* tissue data. The 3-slice model further reduced the through-plane physical needle tip localization error to 2.3±1.1 mm in situations where the imaging plane may be misaligned with the needle. The processing time of the framework using both models was 200 ms per frame. The proposed framework can achieve physical needle localization in real time to support MRI-guided interventions.

INDEX TERMS Interventional MRI, device tracking, needle susceptibility, deep learning, convolutional neural network.

I. INTRODUCTION

The success of minimally invasive image-guided percutaneous procedures, such as targeted biopsy and focal ablation, depends on intra-procedural imaging to visualize tissues and devices (i.e., needles) simultaneously for guidance and confirmation [1], [2]. Magnetic resonance imaging (MRI) has multiple advantages for intra-procedural imaging, including high soft tissue contrast, flexible plane orientations, and no ionizing radiation [3]–[6]. In addition, recent

developments of robotic needle control based on real-time MRI have potential to achieve dynamic needle placement with high accuracy [7]. Automatic, accurate, and rapid needle localization will be required for needle adjustment under both intra-procedural and real-time MRI guidance.

Automatic needle localization based on the passive needle feature is challenged by variations of the needle susceptibility-induced signal void feature due to different situations [8], [9]. Recently, we have developed a deep learning-based method to segment and localize the needle feature on MRI at the instance level using mask region-based convolutional neural networks (Mask R-CNN) [10].

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

R-CNN directs the strong classification power of CNN to certain regions instead of each pixel. An additional mask branch is integrated with the R-CNN to form an end-to-end instance segmentation network [11]. Our model trained with manual annotations of the needle feature achieved rapid (60 ms/frame) and accurate (median tip error around single pixel) needle feature detection and segmentation on different types of MR images. Furthermore, the processing is in real-time since no additional ensemble method, as adopted in other pixel-based CNN segmentation methods [12], [13], is required. Although these emerging deep learning-based techniques show promise for accurate detection and segmentation of the needle feature on MRI, 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 the needle and surrounding tissue causes magnetic field perturbation and MR signal dephasing [8]. With MR-compatible needle materials, such as titanium alloys, the needle susceptibility and geometry usually lead to a long tubular signal void feature on MR images. This needle signal void feature can have an irregular shape at the tip and the axis can be shifted from the physical needle axis [9], [12], [14]–[16]. Therefore, even if the image plane is perfectly aligned with needle, there may be discrepancies between the needle feature and physical needle [9], [14], [17]. Previous studies [15], [17] have reported that this discrepancy can reach 5-10 mm and depends on the MRI sequence type and parameters, the needle material, and the needle's orientation relative to the B_0 field. For reference, clinically relevant tumors for image-guided interventions may have a diameter of 5-10 mm [18]. Therefore, only localizing the needle feature to monitor the physical needle position during the procedure may cause substantial errors in needle targeting.

There are some approaches that have potential to overcome this limitation. One approach is to reduce or correct the distortion of the signal void feature vs. the physical object with multispectral MR imaging. For example, slice-encoding metal artifact correction (SEMAC) can minimize the average needle tip error ($\sim$0.4 mm) with improved time efficiency using compressed sensing (CS) reconstruction [19]. However, the combined acquisition and reconstruction time of CS-SEMAC ($\sim$30 sec) is still not appropriate for immediate updates of the needle position for feedback during procedures. Another approach is to reconstruct the precise physical object shape by forming an inverse problem based on a set of acquired MRI signals. The forward modeling of the needle susceptibility-induced signal void has been reported for different sequence parameters and needle geometry [9], [20]. However, due to the ill-posed nature of the inversion problem, multi-orientation sampling and iterative computation similar to strategies for quantitative susceptibility mapping may be required [21], which are not practical for time-sensitive interventional procedures.

Supervised deep learning using CNNs is a potential approach to rapidly and accurately calculate solutions to ill-posed inversion problems involving magnetic susceptibility. For example, DeepQSM [22] and QSMNet [23] both use pixel-level semantic models based on U-Net [24] to solve ill-posed field-to-source inversion problems and reconstruct quantitative tissue susceptibility maps from single-orientation MRI phase data with rapid inference time. Hence physical needle localization based on the needle feature may potentially be achieved using a deep learning-based approach. These previous pixel-based CNN models aimed to solve for the tissue susceptibility map over the entire field of view (FOV) based on the phase map, but this may not be suitable for the physical needle localization problem, which requires local information about the device. In addition, pixel-level semantic methods could be sensitive to false detection of small objects (e.g., a needle segment in a full FOV image). Therefore, a deep learning-based framework using an instance-level method such as Mask R-CNN can have advantages for physical needle localization.

In MRI-guided interventions, single-slice and multi-slice imaging are used to inform different aspects of the procedure. Imaging plane orientation selection for intra-procedural single-slice and multi-slice imaging is typically performed using information from a multi-slice planning scan in order to ensure that the needle feature is aligned with the slices with minimal through-plane deviation. Single-slice MRI is usually performed to achieve dynamic needle guidance with fast imaging speed [3], [25]. However, through-plane misalignment between the image plane and needle could lower physical needle localization accuracy using single-slice MRI. Sequential multi-slice and orthogonal plane acquisitions have the potential to assist interpretation of this misalignment by physicians [26], [27]. In particular, 3-parallel-slice imaging is a recommended choice from previous studies [26], [28] and our clinical interventional radiologist colleagues.

The main objective of this study was to develop and test an instance-level deep learning framework using Mask R-CNN to automatically and rapidly localize the physical needle position from the passive needle features on MRI. This study first established a single-slice physical needle localization model based on Mask R-CNN by assuming the physical needle was perfectly aligned with the MRI scan plane and focused on addressing in-plane discrepancies between the needle feature and the physical needle. In addition, this study established a 3-slice physical needle localization model based on Mask R-CNN to estimate the physical needle position in the in-plane and through-plane dimensions for 3D localization. Preliminary versions of this work were previously published as a conference abstract [29] and a chapter in the author's Ph.D. dissertation [30].

II. METHODS
A. OVERVIEW OF THE PHYSICAL NEEDLE LOCALIZATION FRAMEWORK
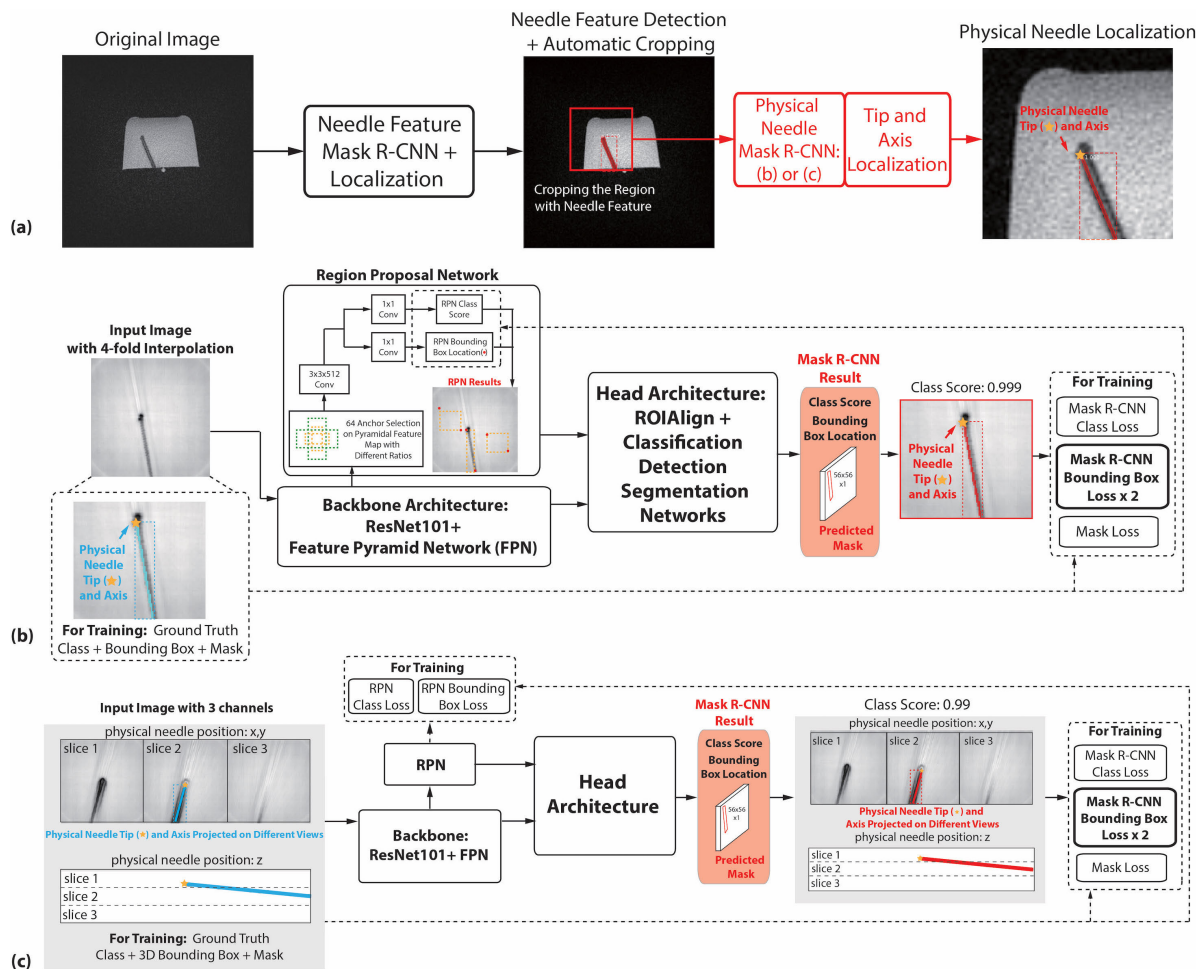We developed a new physical needle localization framework consisting of two Mask R-CNN stages (Figure 1), with

FIGURE 1. (a) The overall physical needle localization framework consisted of two stages of Mask R-CNN. In both stages, the region proposal network (RPN) was initialized with 64 anchors for candidate region creation. The first "needle feature" Mask R-CNN detected and segmented the needle feature on an MR image with field-of-view (FOV) of 300×300 mm^2. Next, the single-slice image or 3-slice images with needle features were automatically cropped to an in-plane FOV of 75×75 mm^2 centered at the detected needle feature tip and used as the input to the second "physical needle" Mask R-CNN, with options (b) and (c), which detected the 2D or 3D physical needle tip and axis. (b) The single-slice physical needle Mask R-CNN was trained using physics-based simulated data of single-slice MRI. Ground truth labels for training were structured as a 2D bounding box with corners that defined the physical needle tip location and axis orientation. Dashed lines indicate data labels and computation steps for training. (c) The 3-slice physical needle Mask R-CNN was trained using physics-based simulated data of 3 adjacent parallel slices of MRI. Ground truth labels for training were structured as a 3D bounding box with corners that defined the physical needle tip location and axis orientation.

the software implementation based on Keras and Tensorflow [31]. For the first stage, we used a "needle feature" Mask R-CNN that we previously trained to detect and segment the needle feature on the input MR image [10]. Based on the results of the first stage, the image was automatically cropped to a patch centered on the needle feature. For the second stage, we trained two separate "physical needle" Mask R-CNN models: 1. A single-slice physical needle Mask R-CNN that takes a needle feature patch from a single slice as input to localize the in-plane 2D physical needle tip and axis (Figure 1b), and 2. A 3-slice physical needle Mask R-CNN to localize 3D physical needle position (in-plane and through-plane) on three adjacent and parallel slices (Figure 1c). Single-slice and multi-slice MRI simulations and experiments were performed to train, validate, and test the proposed framework for physical needle localization.

Ex vivo tissue phantom data were acquired for testing, since the tissue features resemble features on *in vivo* MRI and the phantom setup allowed us to measure the physical needle position under controlled conditions to serve as the reference.

B. MRI-GUIDED NEEDLE INSERTION EXPERIMENTS
We used a golden-angle (GA) ordered radial spoiled gradient-echo (GRE) sequence [32] for real-time 3T MRI-guided needle (20 gauge, 15 cm, Cook Medical) insertion in phantoms. Three sets of imaging parameters (echo time [TE] = 1.9 ms, readout bandwidth [BW] = 888 Hz/pixel; TE = 3 ms, BW = 888 Hz/pixel; TE = 2.8 ms, BW = 300 Hz/pixel) and different needle orientations were used to create variations in the passive needle feature on MRI. To achieve sufficient image quality and signal-to-noise ratio (SNR) for needle visualization, a temporal resolution of around 1 sec/frame was used for

TABLE 1. Imaging parameters for real-time radial radiofrequency (RF)-spoiled gradient-echo (GRE) MRI of the gel phantom and *ex vivo* tissue phantom at 3T. Three combinations of repetition time (TR), echo time (TE), and bandwidth (BW) were used to acquire images with different tissue contrast and needle signal void feature. Similar imaging parameters were used to generate simulated images that form the network training dataset. 3-parallel-slice input used the same imaging parameters and had temporal resolution of 2.4-3.6 sec/image.

Real-Time Radial RF-Spoiled GRE 3T MRI			
TR, TE, BW	4.16 ms, 1.9 ms, 888 Hz/pixel	5.26 ms, 3 ms, 888 Hz/pixel	6.3 ms, 2.8 ms, 300 Hz/pixel
Field of view	300 mm x 300 mm	**Flip angle**	12°
Matrix size	256 x 256	**Dynamic temporal resolution (reconstruction window length)**	0.8-1.2 sec / image (200 radial spokes)
Slice thickness	5 mm	**Orientation**	Oblique coronal

single-slice real-time MRI. The imaging parameters are listed in Table 1. Images were reconstructed using non-Cartesian conjugate gradient sensitivity encoding (SENSE) with reconstruction window size of 200 radial spokes [32], [33].

We performed needle insertion experiments in gel and *ex vivo* tissue phantoms (Figure 2a). The *ex vivo* tissue (pork shoulder or beef chuck steak) was submerged underneath gelatin to create a flat surface on the top. An MRI-visible fiducial marker (MR-SPOT, Beekley Medical, Connecticut, USA) was affixed to the surface to define the needle entry point (Figure 2b). To directly control and confirm the physical needle orientation and insertion depth during experiments, we used an MRI-compatible master-slave needle actuator system [34]. We defined the orientation of the needle in terms of its rotation angle (θ) in one plane and tilting angle (α) in the orthogonal plane (Figure 2c). A linear guide on the actuator ensured a straight needle insertion path and a fixed length of the needle (50 mm) was inserted so that the physical needle tip can be measured in 2D image coordinates based on the position of the fiducial marker. These measurements were taken as the references of the physical needle tip position and axis orientation (Figure 2d) for single-slice physical needle Mask R-CNN. The range of needle axis orientation is reported in Table 2.

In addition, a similar set-up was used to acquire the reference physical needle position for 3-slice physical needle Mask R-CNN. One smaller MRI-visible fiducial marker was affixed on the surface to mark the needle entry point and another longer fiducial marker was attached on the needle shaft. 3D high-resolution T_1-weighted gradient echo images with isotropic voxel size of $(0.6 \text{ mm})^3$ were acquired to measure the physical needle position based on the pre-determined position of the entry point relative to the small fiducials and a fixed needle insertion length (60 mm) (Figure 3b-d). The reference physical needle tip positions were measured in the RAS (right/left, anterior/posterior, superior/inferior)
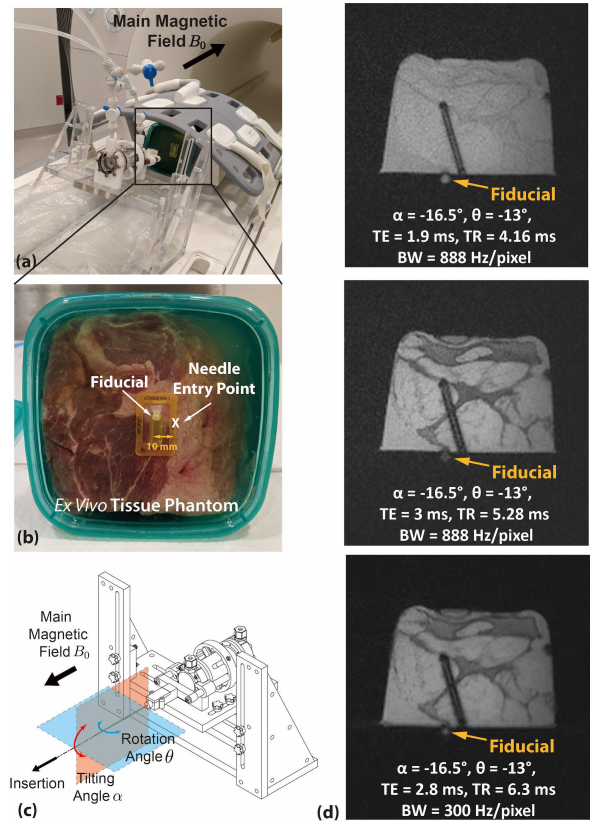

FIGURE 2. (a) Experimental setup in the 3T MRI scanner. (b) An MRI-visible fiducial marker was affixed to the phantom to measure the physical needle position in the *ex vivo* tissue phantoms. (c) In phantom MRI experiments, the needle orientation was varied in terms of its rotation (θ) and tilting angle (α) relative to B_0. A needle actuator system was used to insert the needle into a phantom by 50 mm without bending. The ranges of α and θ are reported in Table 2. (d) Example *ex vivo* tissue phantom MR images ($\alpha = -16.5°$ and $\theta = -13°$) with three different sets of imaging parameters. All phantom MR images were reconstructed using conjugate gradient SENSE. TE: echo time. TR: repetition time. BW: readout bandwidth.

TABLE 2. Physics-based simulated datasets used for training of single-slice physical needle Mask R-CNN (top row). Datasets used for fine-tuning and testing of the overall physical needle localization framework with two Mask R-CNN stages (middle and bottom). θ: needle rotation angle. α: needle tilting angle. N/A: not applicable. *Data augmentation included rescaling, translation, and additive Gaussian noise.

	Data size (images)		Needle orientation		
	Original	After 5-fold augmentation*	θ	α	Step
Training dataset	741	3705	-30° to 30°	-90° to 0°	5°
Fine-tuning dataset	741	3705	-30° to 30°	-90° to 0°	5°
Testing dataset	186	N/A	-34° to 41°	-92° to -7°	N/A

coordinate system using open-source software (3D Slicer) [35], The range of needle axis orientation is reported in Table 3.

TABLE 3. Physics-based simulation dataset for training of 3-slice physical needle Mask R-CNN and *ex vivo* tissue datasets used for testing the overall physical needle localization framework. θ: needle rotation. α: needle tilting angle relative to B₀ field. η: needle-to-slice orientation difference. h: pivot point. N/A: not applicable. * Data augmentation included rescaling, translation, and additive Gaussian noise.

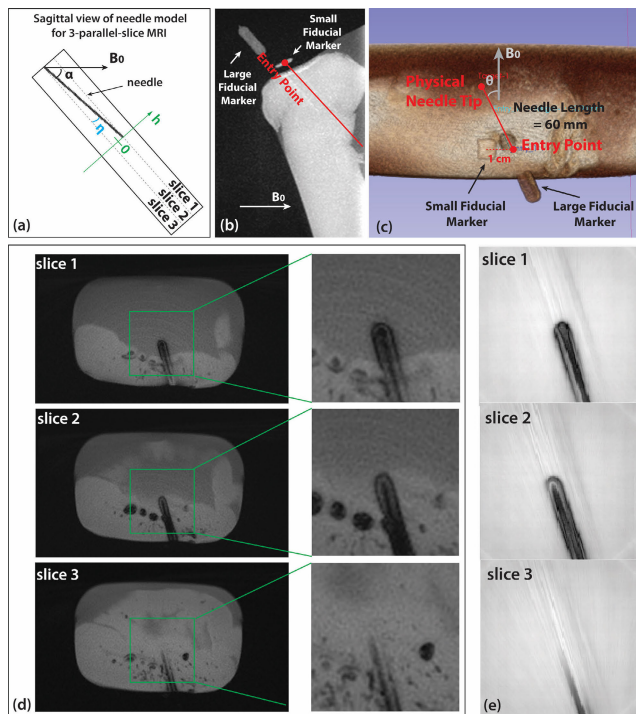|	Data size (sets of 3-slice images)		Needle orientation				
	Original	After 5-fold augmentation*	θ	α	η	h	Step
Training dataset	2160	10800	-30° to 30°	-90° to 0°	6° to -6°	-5 mm to 5 mm	5°
Testing dataset	150	N/A	-22° to 32°	-72° to -15°	7° to -5.5°	-3.8 mm to 5 mm	N/A

FIGURE 3. (a) An example diagram showing the sagittal view of an inserted needle misaligned with 3 parallel oblique coronal imaging planes. Needle tilting angle (α) is −49.1°, needle-to-slice orientation difference (η) is 5.5° and pivot point (h) is 2.5 mm. (b) maximum intensity projection of high-resolution 3D MRI in the sagittal view showed the fiducial markers used to locate the needle tilting angle (α) relative to the main magnetic field B₀ and entry point. (c) 3D rendered model of the phantom and the fiducial markers in a 3D environment. Physical needle tip is determined based on fixed needle insertion length (60 mm) and entry point that was marked at the centerline of the small fiducial marker, which was 1 cm from its center point. Needle rotation angle (θ) is −17.9°. (d) Corresponding 3-parallel-slice MR images in the actual needle insertion experiment in a phantom. (e) Simulated passive needle features showed similar feature pattern and distribution on different imaging planes.

C. NEEDLE SUSCEPTIBILITY SIMULATION AND CALIBRATION

Training the physical needle Mask R-CNN requires a substantial amount of data with reliable reference needle position information. However, collecting data from the MRI experiments described in the previous section is expensive, time-consuming, and subject to experimental uncertainties. Therefore, we performed MRI physics-based simulations of

the needle susceptibility effects to generate a large set of training images that accurately depict the needle feature with respect to actual intra-procedural MRI.

We implemented the Fourier-based off-resonance artifact simulation in the steady state (FORECAST) method to calculate the susceptibility effects in steady-state GRE MRI [9]. The field inhomogeneity or field shift $\Delta B_0 (x, y, z)$ was calculated as a function of different needle orientations and needle materials with different magnetic resonance properties using a first order perturbation approach to Maxwell's equations, combined with the Fourier transformation technique [36]. In the original FORECAST method, a thin slice with the desired FOV and slice thickness was modeled in 3D space, with the third dimension of ΔB_0 set to be parallel to B_0, which does not capture the tilting angle of the needle. To simulate the needle with a tilting angle, which is a more realistic scenario in interventional procedures, we created an expanded 3D model (Figure 4). Specifically, $\Delta B_0 (x, y, z)$ was calculated and re-sliced to an oblique volume parallel to the needle with certain excitation slice or slab thickness. A linear interpolation step was performed to assign the ΔB_0 to each pixel of the model with the original pixel dimensions. In addition, a non-uniform fast Fourier transform (NUFFT) was applied for the GA ordered radial sampling trajectory during the simulations [37]. The overall k-space signal (s) model of the needle susceptibility-induced effects on the discrete isochromatic grid with proton density $\rho' (x, y, z)$ was:

$$s\left(k_x, k_y, k_z\right) = \sum_x \sum_y \sum_z \rho'\left(x, y, z\right) e^{-i2\pi \gamma \Delta B_0(x,y,z)t' - k_z t'}$$
$$\times e^{-\frac{t'}{T_2}} e^{-i2\pi\left(k_x x + k_y y\right)} \quad (1)$$

where γ is the gyromagnetic ratio and t' is the time after radiofrequency (RF) excitation. Note that the T_2^* of the signal was decomposed into T_2 (e.g., 50 ms for muscle) and the field inhomogeneity caused by the needle susceptibility effects. Finally, an inverse NUFFT was applied to the simulated k-space data to reconstruct the image, which contains the needle signal void feature.

For these simulations, we assumed that the needle material is stiff enough [38] and there was no deflection close to the tip. Therefore, we modeled the needle as
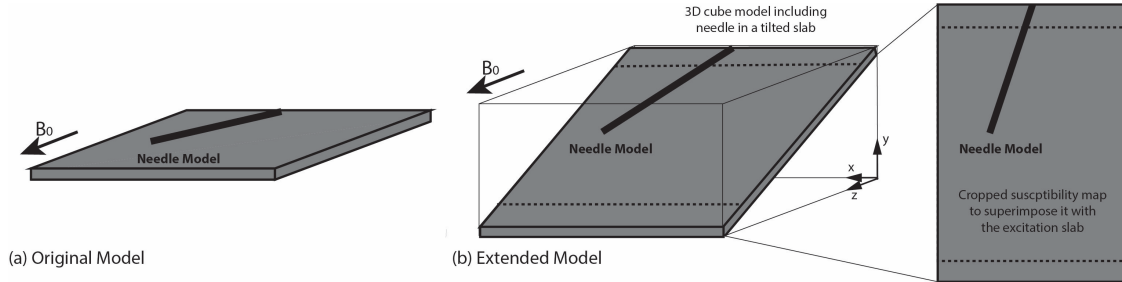
FIGURE 4. (a) In the original FORECAST method, the susceptibility map can only be calculated for an imaging slice/slab that has one direction aligned with the main magnetic field B_0. (b) To calculate the susceptibility map for a needle model that is tilted with respect to B_0, we created an expanded 3D model. The susceptibility map in this 3D model was interpolated to the coordinates and model elements in the excited imaging slice/slab, which contained the needle with a certain tilting angle. The gray plane represents a single slice or a thin slab of multiple parallel slices excited for imaging. At the end, the susceptibility map was cropped to match the size of the imaging slice/slab.

a cylindrical rod with diameter of 0.9 mm (20 gauge) at different rotation (θ) and tilting (α) angles in 3D space similar to the actual experimental setup (see previous section). The range of θ ($-30°$ to $30°$) and α ($0°$ to $-90°$) of the needle was set according to actual reports of needle placement in abdominal percutaneous interventions by other groups [17], [39], [40] and our clinical colleagues.

To ensure that the simulations matched the conditions of the needle used in experiments, we used actual experimental data from a gel phantom with different needle orientations and imaging parameters (Table 1) to calibrate the susceptibility value of the needle material. The Euclidean distance between the physical needle tip and the needle feature tip (dxy in mm, Figure 5) was calculated for simulated data and gel phantom experimental data. The susceptibility value that minimized the average dxy was identified and used in simulations to generate training data.

D. SINGLE-SLICE PHYSICAL NEEDLE MASK R-CNN TRAINING, FINE TUNING AND TESTING

The input to the single-slice physical needle Mask R-CNN is an image patch centered on the needle feature and surrounding tissue, obtained from the needle feature Mask R-CNN output. The training dataset for the single-slice physical needle Mask R-CNN consisted of simulated images with the same size as the expected input patches (75×75 mm^2 FOV with 256×256 matrix size) and the needle feature tip in the center of the patch (Figure 1). The input 2D gray-scale image patches were normalized to pixel intensities of 0-255 and stacked into 3 color channels to match the dimensions of the pre-trained needle feature Mask R-CNN model. We used the output bounding box corners to define the physical needle tip (e.g., at the top left corner) and axis (e.g., the line connecting the top left and bottom right corners) since the diameter of the actual physical needle is less than the pixel size. Since the needle tip location and axis orientation were more important than the needle segmentation mask in the physical needle Mask R-CNN stage, we weighted the bounding box loss to twice that of the losses in other branches during training.
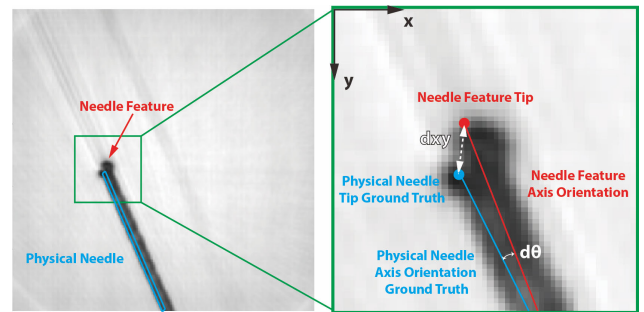

FIGURE 5. The discrepancy between the needle feature position and the actual physical needle position was quantified in terms of the Euclidean distance between the feature tip and the physical needle tip (dxy), and the absolute difference between the feature axis and the physical needle axis orientation (dθ), both in image coordinates in 2D space. The accuracy of physical needle localization using the proposed framework with single-slice physical needle Mask R-CNN was also quantified in terms of dxy and dθ.

We first generated simulated data for single-slice physical needle Mask R-CNN training. A set of 741 simulated images with the same parameters as phantom experiments (θ from $-30°$ to $30°$, and α from $0°$ to $-90°$) was created for training (Table 1). Five-fold data augmentation was performed by rescaling, translation, and adding Gaussian noise to form a training dataset with 3705 images. A small rescaling factor (95% or 105%) was used to minimize changes in pixel size and maintain validity of the simulated physical needle model. The training hyperparameters were based on our previous work on needle feature segmentation [10] and the number of epochs (480 epochs) was increased to accommodate the larger data size. In a preliminary study [29], model validation using simulated data confirmed that the majority of physical needle tip localization error was less than the model element size (0.3 mm). This close agreement with the ground truth corroborates the capability of the physics-based model to predict the in-plane physical needle location in different situations and there is no need to further expand the training dataset, which considerably reduced the training time.

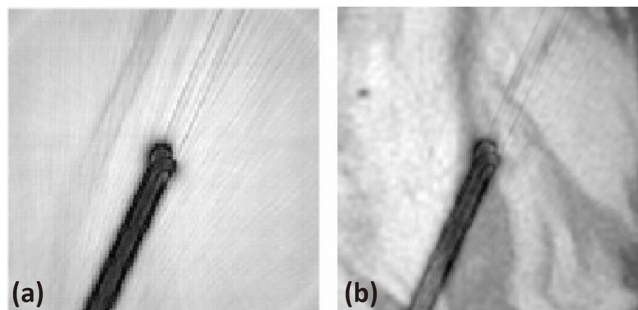However, interventional MR images acquired during actual procedural guidance have more complex backgrounds

FIGURE 6. (a) An example training image generated from MR physics simulation. (b) An example image for physical needle Mask R-CNN model fine-tuning, which was created by superimposing the simulated needle on a patch of tissue background from *ex vivo* MRI.

compared to the simulated images. In addition, certain types of tissue with off-resonance effects (e.g., fat) may also generate signal voids (e.g., fat-water signal cancellation) that occlude the needle feature. These effects might degrade the accuracy of our framework. Therefore, after training with simulated data as previously described, we performed fine-tuning of the physical needle Mask R-CNN by using an additional training dataset with enriched variations of the background. To do this, we acquired MR images of *ex vivo* tissue in different slices without a needle. Patches were randomly cropped from these *ex vivo* tissue images and superimposed with the simulated needle images (741 images) (Figure 6), followed by similar data augmentation to increase the size of the fine-tuning dataset (3705 images) (Table 2).

We tested the physical needle localization accuracy of the proposed single-slice two-stage framework using *ex vivo* tissue phantoms, which have realistic image characteristics that are representative of interventional MRI. 186 images were collected from experiments to form a testing dataset with 62 different needle orientations (Table 2). We compared the accuracy of physical needle localization using the single-slice physical needle Mask R-CNN models without and with fine-tuning. The Euclidean distance between the estimated physical needle tip position from the two-stage framework and the measured reference physical needle position (dxy in mm) was calculated based on image coordinates (Figure 5). The absolute difference between the needle axis orientations ($d\theta$ in degrees) was also computed to evaluate the localization accuracy.

E. THREE-SLICE PHYSICAL NEEDLE MASK R-CNN TRAINING AND TESTING

In addition to the single-slice strategy, we also developed a Mask R-CNN model for physical needle localization using 3 parallel slices. The bounding box output was increased from 4 to 6 dimensions to capture the 3D coordinates of the physical needle's two ends (tip and entry point). The input to this network was acquired in the same approach as the single-slice network, except that it contains three image patches, including needle features from three adjacent

parallel slices. The three patches were normalized to pixel intensities of 0-255 and stacked into three color channels of the network input. Then we applied the same simulation method to generate the training dataset. The thickness of the slab in the 3D model (Figure 4) was expanded from 5 mm to 15 mm to emulate three parallel imaging slices without any gap. Misalignment between the needle model and 3D acquisition slab was characterized by two additional geometric parameters: needle-to-slice orientation difference (η) and pivot point (h) within the imaging slab (Figure 3a).

To determine the range of the additional parameters for simulations, needle obliqueness with respect to the imaging plane was analyzed in clinical MRI-guided prostate biopsy datasets (326 images from 125 cases). We found that the misalignment of the image plane with the needle feature had a median of 3.1° with interquartile range of 2.6° [10]. Around 92% of the cases had misalignment between needle feature and imaging plane of <6°. The passive needle features were contained within 1 to 3 slices (slice thickness of ~5 mm) for each imaging dataset based on the relative position between the imaging plane and needle feature axis. Based on these typical ranges of needle angles, we can assume that the tilted needle feature would be contained within 3 parallel slices with 5-mm slice thickness. Therefore, we considered 35 representative parameter combinations of $\eta = [-6°, -4°, -2°, 0°, 2°, 4°, 6°]$ and h = [5, 2.5, 0, -2.5, -5] mm. The physical needle reference was defined with a 3D bounding box and the third-dimension coordinate can be visualized from the side view of the 3D acquisition slab (Figure 3a).

We trained the 3-slice physical needle Mask R-CNN for needle localization using a set of 2160 simulated images with the same parameters as phantom experiments (θ from $-30°$ to $30°$, and α from $0°$ to $-90°$) (Table 3). 10 random choices of different combinations of η and h were selected for each imaging plane orientation. *Ex vivo* tissue images were superimposed with these simulation images and 5-fold data augmentation was performed to form a training dataset with 10800 images. The epoch number was increased to 480 to accommodate the larger training dataset. The network parameters were initialized with random weights instead of using any pretrained single-slice physical needle Mask R-CNN model because of the structural differences of the 3-slice network and different feature encoding of channel-wise information from input images.

We tested the needle localization accuracy of the proposed two-stage framework in *ex vivo* tissue phantoms using 3-parallel-slice MRI. 150 images were collected from experiments to form a testing dataset with 25 different needle orientations (Table 3) that were misaligned with the imaging plane orientation. At each needle position, three sets of imaging parameters with two different η were used to acquire the imaging dataset for testing. We compared the accuracy of physical needle localization using the single-slice physical needle Mask R-CNN and 3-slice physical needle Mask R-CNN. One of the 3-parallel-slice images with the complete needle feature was selected as

the input for the single-slice physical needle Mask R-CNN. The error metrics were all measured in 3D (tip localization error: dxyz; 3D orientation difference: dϕ). dxyz was calculated as the Euclidean distance between the estimated and reference physical needle tip positions. dϕ was calculated as the orientation difference (angular separation) between the estimated and reference physical needle axes in 3D space.

F. PHYSICAL NEEDLE LOCALIZATION ERROR ANALYSIS

For tip localization error (dxy or dxyz), we considered a 2.5-mm threshold since clinically relevant tumors for interventions may have diameters of 5-10 mm [18]. In addition, the needle tip and axis orientation discrepancy between the needle feature and physical needle were manually measured by a trained researcher to serve as a baseline to assess the improvement using the proposed Mask R-CNN framework. The physical needle localization results using different models and the error from the needle feature discrepancy were first compared using a non-parametric analysis of variance (ANOVA) (i.e., Kruskal-Wallis test). If significant differences were detected by the Kruskal-Wallis test, the medians of the results were compared in a pair-wise fashion using the non-parametric Wilcox signed rank test. In addition, the variances of the results were compared using the Brown-Forsythe test. We considered differences with p <0.05 to be significant.

III. RESULTS

A. NEEDLE SUSCEPTIBILITY CALIBRATION

Simulated MR images that contained the needle feature with different rotation and tilting angles were compared with the MR images from actual MRI-guided needle insertion experiments using a gel phantom (Figure 7). The characteristics of the needle feature from simulations closely matched the needle feature on experimental images, with one difference being the additional noise seen on experimental images. The physical needle position in the coordinates of the experimental MR images were determined based on the fiducial marker position. The spatial relationship (discrepancy) between the physical needle and the needle feature were almost identical for simulated and experimental images. During the calibration process, we found that a needle susceptibility value of 190 ppm (corresponding to titanium) achieved close agreement between the simulations and experiments for 7 different rotation angles (Figure 7). Therefore, we used this calibrated susceptibility value for subsequent simulations. Example 3-parallel-slice simulation images using this susceptibility value also generated needle features that are in close agreement with the MR images from actual experiments (Figure 3d-e).

B. SINGLE-SLICE PHYSICAL NEEDLE MASK R-CNN TRAINING AND TESTING

We trained the single-slice physical needle Mask R-CNN using a batch size of 8 on two graphics cards (NVIDIA GTX 1080Ti). Generation of the simulated data for
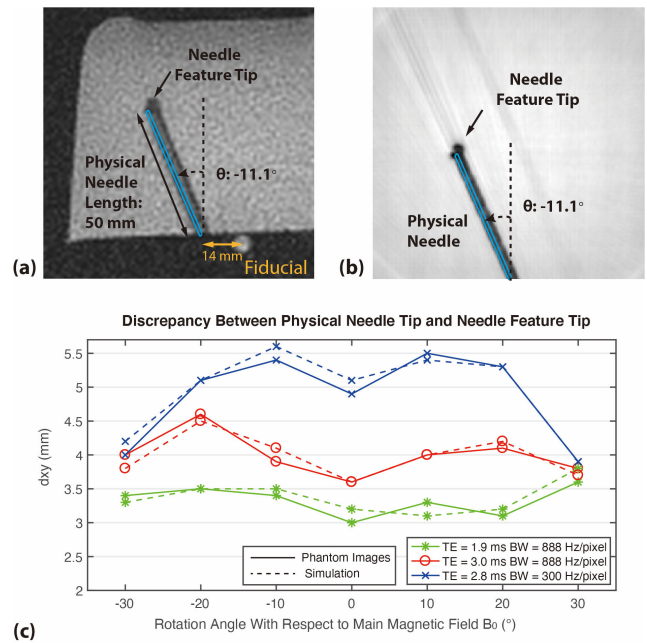

FIGURE 7. (a) An example real-time 3T single slice MR image with needle tilting angle (α) of 11.1° and rotation angle (θ) of −24° with respect to B_0, TE = 1.9 ms, and bandwidth (BW) = 888 Hz/pixel. (b) Example simulated image using a $300 \times 300 \times 5$ mm^3 field-of-view (FOV) with $1024 \times 1024 \times 17$ model elements, 256 radial readout points, and a 20-gauge needle with the same orientation as in (a). The images were reconstructed using non-uniform fast Fourier transform (NUFFT) and then cropped to a FOV of 75×75 mm^2 with image matrix size of 64×64. (c) Calibration results of needle susceptibility (190 ppm) showing dxy for seven different needle rotation angles.

training took 10 hours. Training with simulated data took about 12 hours and the fine-tuning step (simulated needle features combined with tissue image patches) took another 12 hours.

Representative results for physical needle localization in *ex vivo* tissue testing data using the fine-tuned single-slice Mask R-CNN model are shown in Figure 8a-c. The processing time of the whole framework was about 200 ms per frame. These examples showed not only the needle localization accuracy of the framework, but also the improvement of the fine-tuned single-slice physical needle Mask R-CNN model compared to the non-fine-tuned model. Figure 8d shows an example where the fine-tuned model successfully detected the physical needle, but the model without fine-tuning failed to detect the needle. Figure 8e shows an example with accurate physical needle tip localization and axis orientation estimation using the fine-tuned model. In contrast, the tip localization results for this case had large errors (dxy = 5.76 mm) using the model without fine-tuning.

The overall testing results for the two-stage framework using the single-slice physical needle Mask R-CNN models without and with fine-tuning are shown in Table 4. The framework using the non-fine-tuned model achieved median dxy and dθ of 0.94 mm and 0.64°, respectively, while the framework using the fine-tuned model reduced the median dxy and dθ to 0.81 mm and 0.63°, respectively.
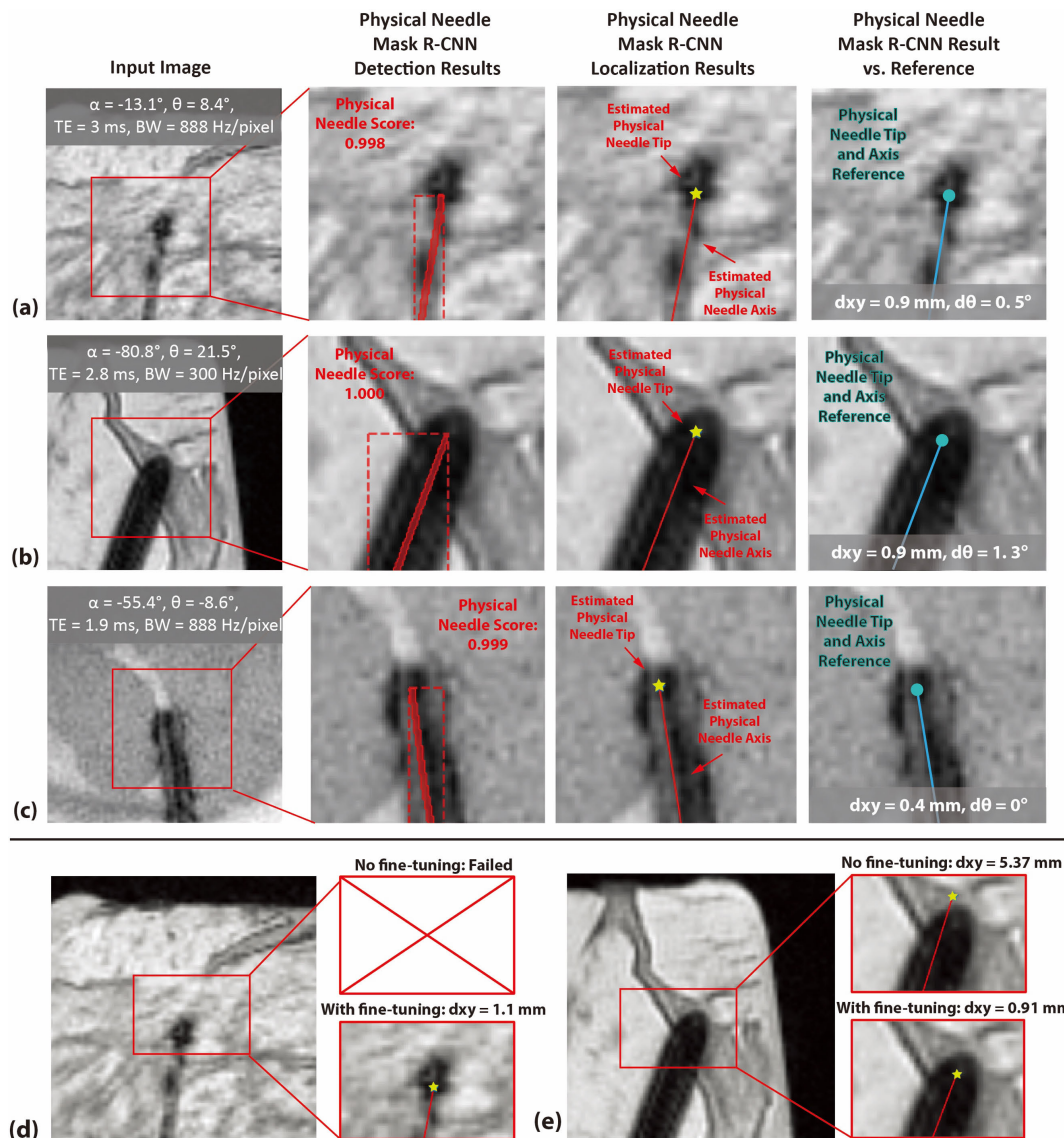
FIGURE 8. (a-c) Example single-slice physical needle Mask R-CNN localization results in three *ex vivo* tissue phantom testing images with different imaging parameters and different needle α and θ. The needle tip location (dxy) and axis orientation (dθ) differences compared to the reference (measurement during experiments) are reported in each example. (d) Example of much larger physical needle localization error using the model without fine-tuning compared to the model with fine-tuning. (e) Example of needle detection failure using the model without fine-tuning and success using the model with fine-tuning.

In addition, the distributions of these results are summarized and compared using violin plots (Figure 9). Figures 9a-b show dxy and dθ using the proposed framework (without and with fine-tuning) and the discrepancy between the needle feature tip/axis orientation and physical needle tip location/axis orientation. There were differences in dxy (p = 8×10^{-78}) and dθ (p = 2.7×10^{-11}) among these three sets of results. The pair-wise comparison among 3 sets of results indicated that the median and variance of dxy were both significantly reduced by using the fine-tuned model. There were significant differences in the median and variance of dθ between the non-fine-tuned and fine-tuned models

compared to the discrepancy between needle feature and physical needle axis orientation, but there were no significant differences between the non-fine-tuned and fine-tuned models. More details about significance and p value are shown in Figure 9.

C. THREE-SLICE PHYSICAL NEEDLE MASK R-CNN TRAINING AND TESTING

Overall training time for the 3-slice physical needle Mask R-CNN model was 48 hours. Representative results of 3D physical needle localization projected on 2D image slices and an orthogonal side view of the estimated physical

TABLE 4. Physical needle localization accuracy using the proposed models in the testing datasets (see Table 2). Needle tip position errors (*dxy* or *dxyz*) and absolute needle axis orientation differences (*dθ* or *dφ*) are reported. SD: standard deviation. IQR: interquartile range. Success: the physical needle Mask R-CNN detected the physical needle in the image. *Pixel size of the acquired image. #2.5 mm threshold considering clinically relevant tumor sizes of 5 mm diameter.

|			Mean	SD	Median	IQR	Max	dxy < 1.17 mm* or dθ < 1° (% of images)	Success rate
Ex vivo tissue phantom single-slice testing dataset (186 images)	No fine-tuning	dxy	1.24 mm	0.97 mm	0.94 mm	1 mm	5.7 mm	62.4%	185/186 (99.5%)
		dθ	0.79°	0.7°	0.64°	0.7°	4.47°	73.6%	
	With fine-tuning	dxy	0.96 mm	0.69 mm	0.81 mm	0.65 mm	4.3 mm	74.7%	186/186 (100%)
		dθ	0.75°	0.7°	0.63°	0.63°	3.58°	76.4%	
			Mean	SD	Median	IQR	Max	dxyz < 2.5 mm# or dφ < 1° (% of images)	Success rate
Ex vivo tissue phantom 3-slice testing dataset (150 sets of images)	Single-slice model	dxyz	3.1 mm	1.4 mm	2.9 mm	1.5 mm	10.2 mm	35.3%	150/150 (100%)
		dφ	2.4°	1.2°	2.5°	1.2°	6.7°	12%	
	Three-slice model	dxyz	2.3 mm	1.1 mm	2.2 mm	1.2 mm	7.2 mm	57.3%	150/150 (100%)
		dφ	1.4°	0.85°	1.2°	1°	4.7°	52%	
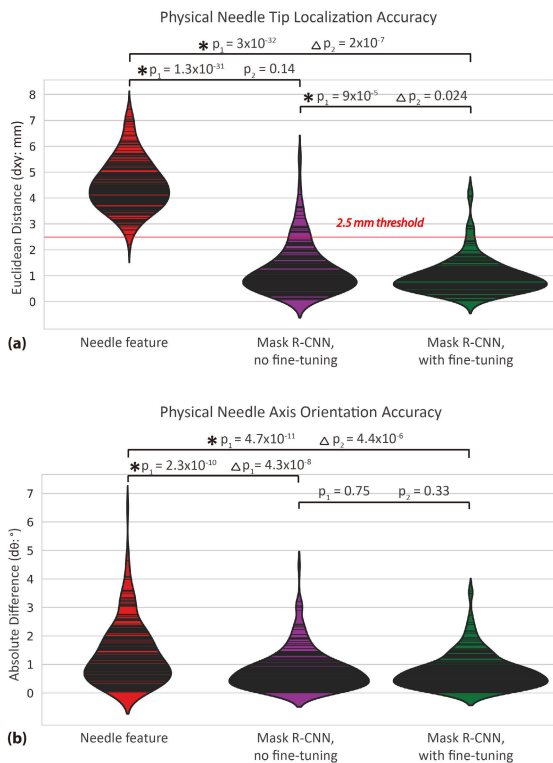
FIGURE 9. Violin plots of single-slice physical needle (a) tip localization and (b) axis orientation results for the ex vivo tissue phantom testing dataset using needle feature localization and the physics-driven Mask R-CNN model without and with the fine-tuning step. In (a), The red line represents the 2.5 mm threshold considering clinically relevant tumor sizes of ≥5 mm diameter. In the pair-wise comparisons, p_1 value of Wilcoxon signed rank test is on the left and p_2 value of Brown-Forsythe test is on the right. * indicates Wilcoxon signed rank test with $p_1 < 0.01$. △ indicates Brown-Forsythe test with $p_2 < 0.05$.

needle using the single-slice and the 3-slice physical needle Mask R-CNN models are shown in Figure 10. These results showed accurate physical needle localization in 3D using the 3-slice model, compared with the reference. Physical needle axis orientation difference was reduced in the through-plane direction using the 3-slice model compared with the single-slice model. The processing time of the whole framework was about 200 ms per set of 3-parallel-slice images, which is suitable for real-time interventional MRI applications. The overall results are summarized in Table 4. The framework using the single-slice model achieved median dxyz and dφ of 2.9 mm and 2.5°, respectively, while the framework using the 3-slice model reduced the median dxyz and dφ to 2.2 mm and 1.2°, respectively.

The distributions of the results from all testing datasets are summarized and compared using violin plots. Figure 11 shows dxyz and dφ using the proposed framework with single-slice and 3-slice models and the discrepancy between the needle feature tip/axis orientation and physical needle tip location/axis orientation. There were differences in both dxyz ($p = 1.1 \times 10^{-59}$) and dφ (2.8×10^{-28}) among these three sets of results. The pair-wise comparison among three sets of results indicated that dxyz and dφ were significantly reduced using the 3-slice model. No significant differences of variance in dxyz were observed from any of these pairs. On the other hand, the variance of dφ was significantly lower by using the 3-slice model. More details about significance and p values are shown in Figure 11.

IV. DISCUSSION

In this study, we developed a new automatic physical needle localization framework for MRI-guided percutaneous interventions. The framework included two Mask R-CNN stages. First, the needle feature Mask R-CNN provided an initial estimate for the needle position. Next, the cropped image patch(es) containing the needle feature was sent to
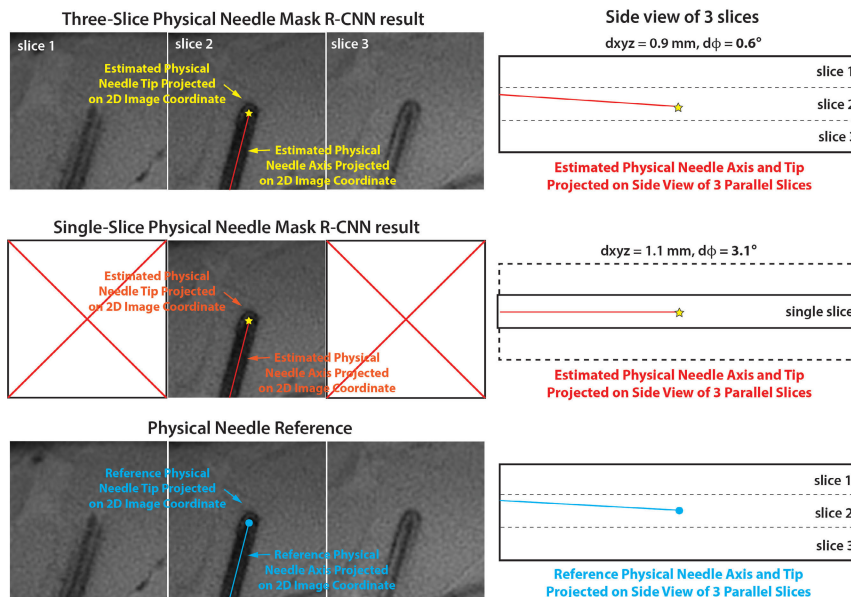
FIGURE 10. Example 3-slice physical needle Mask R-CNN localization results and single-slice physical needle Mask R-CNN localization results projected on 2D in-plane image coordinates and orthogonal side view of 3-slice and single slice MRI with needle tilting angle (α) of $-54.2°$ and rotation angle (θ) of $-19°$ with respect to B$_0$, TE = 1.9 ms, and bandwidth (BW) = 888 Hz/pixel.

a single-slice or 3-slice physical needle Mask R-CNN. The single-slice model takes a single-slice image as input, assuming the imaging plane is already closely aligned with the physical needle. The 3-slice model takes 3 parallel and adjacent slices as input, in which the imaging plane orientation could be misaligned with the physical needle axis. Both models were trained by a substantial set of physics-based simulation images that included realistic needle-induced susceptibility features. The cropped image patch helped to avoid false detection results and maintained the assumption of a rigid needle segment for the input to the physical needle Mask R-CNN models. The reference physical needle position was measured using a fiducial marker and needle actuator, which achieved stable and repeatable needle placement during the experiments [34]. Our validation and testing results demonstrated that the proposed framework with single-slice model accurately and rapidly estimates the in-plane physical needle position using single-slice MRI. The 3-slice model further reduced the through-plane needle localization error due to misaligned imaging plane with physical needle and rapidly estimated the overall 3D physical needle position.

As part of our framework, we developed an image-based needle susceptibility calibration method that compares the discrepancies between the physical needle and needle feature from experimental MRI data with the physics-based simulations in different situations. This calibration step can improve the understanding of the needle feature characteristics under specific conditions and on specific types of MR images. It showed that the discrepancies between needle feature and physical needle varied with different needle orientations and imaging parameters; proper selection of the needle susceptibility ensured the fidelity of the simulated images

for training. By adding noise during data augmentation, the simulated images formed a sufficient dataset to train the physical needle Mask R-CNN while avoiding the need for expensive MRI experiments and manual annotations.

Ex vivo tissue phantom MR images have realistic noise characteristics and also tissue features in the background, which resemble features expected on *in vivo* interventional MRI. The statistical comparisons showed that the physical needle tip localization accuracy was improved by using the fine-tuned singe-slice model trained by fusing simulated needle features with tissue background patches. The percentage of dxy >2.5 mm using the non-fine-tuned model was three times larger than the fine-tuned model. The threshold of 2.5 mm is informative, as clinically relevant tissue targets for minimally invasive interventions have diameters of 5-10 mm or larger [18]. Therefore, the fine-tuned model has potential to reduce mistargeting during MRI-guided interventions, especially for precise maneuvers in smaller targets. Furthermore, the physical needle localization error using the proposed network was lower than the localization error caused by the discrepancy between the physical needle and the needle feature. Overall, our results demonstrate that our proposed deep learning-based framework using single-slice model is an accurate and fast approach (processing time of 200 ms/image) to overcome the discrepancy, thereby achieving accurate physical needle localization on MRI. No additional processing was required except for the two-stage Mask R-CNN, which combined an updated faster R-CNN and a mask branch [11]. The fine-tuning scheme implemented in our framework can be used in future work to adapt it for specific *in vivo* applications (i.e., fine-tuning with additional *in vivo* datasets relevant for an MRI-guided procedure).
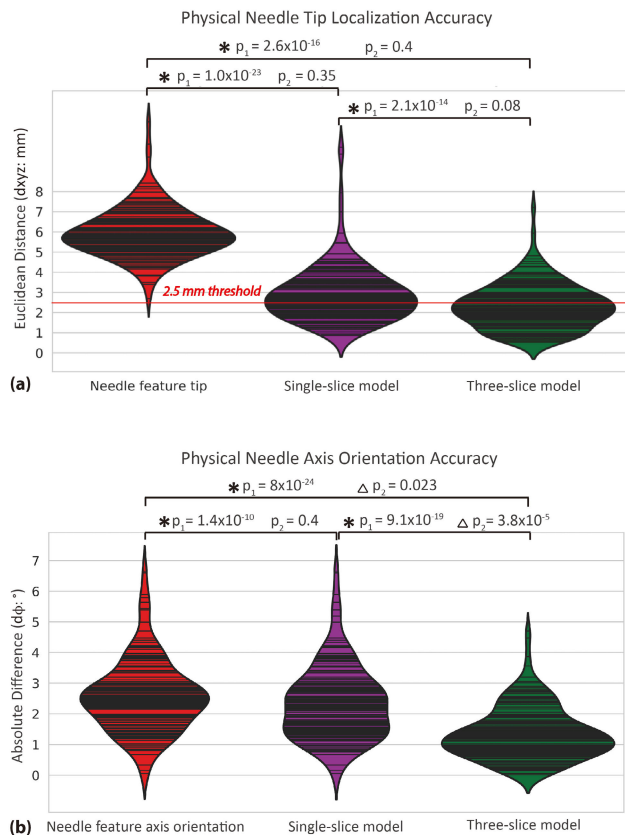
FIGURE 11. Violin plots of physical needle (a) tip localization and (b) axis orientation results for the *ex vivo* tissue phantom testing dataset with certain range of misalignment between the needle axis and image plane orientation (shown in Table 3) using needle feature localization and the physics-driven single-slice and 3-slice Mask R-CNN models. In (a), the red line represents the 2.5 mm threshold considering clinically relevant tumor sizes of ≥5 mm diameter. In the pair-wise comparisons, p_1 value of Wilcoxon signed rank test is on the left and p_2 value of Brown–Forsythe test is on the right. * indicates Wilcoxon signed rank test with $p_1 < 0.01$. △ indicates Brown–Forsythe test with $p_2 < 0.05$.

Our extended FORECAST method for the 3-slice physics-based simulation took about three days and 1 GB storage space to generate the training data. Since the input and output of the 3-slice model and the spatial information encoded from different input channels are different from the single-slice model and other Mask R-CNN models for in-plane object detection and segmentation, no pretrained model was used during the training. *Ex vivo* tissue phantom images for 3-slice model testing considered a specified range of misalignment between the imaging plane and the physical needle. Statistical comparisons of dxyz and $d\phi$ demonstrated that the 3-slice model reduced the through-plane needle localization error compared with the single-slice model. Overall, the median 3D physical needle tip localization error was 2.2 mm and more than 50% of the results were <2.5 mm. This corresponds to 1 to 2 pixels in-plane and is less than the 5-mm slice thickness. While this localization performance is not at the subpixel level, as we have demonstrated for the case when the imaging plane is aligned with the needle, the direct 3D localization accuracy already can be sufficient for certain targeting applications (e.g., targets around 10 mm in diameter).

The single-slice and 3-slice models we have developed in this work can be used in concert to support different steps in the MRI-guided interventional workflow. For example, during procedural setup and adjustment, the 3-slice model can rapidly localize the physical needle position in 3D space and inform alignment of the MRI scan plane(s) with the physical needle, using standard manual adjustments or new automated methods [41]. Once the imaging plane is aligned with the physical needle, the single-slice model can be applied to dynamically localize, track, and confirm the physical needle position with sub-millimeter accuracy. This strategy for closed-loop confirmation of the physical needle position has the potential to improve physicians' confidence in assessing and ensuring procedural accuracy.

There are limitations in this study. Firstly, we only tested the GRE sequence with 3 different sets of imaging parameters that we often use in our research work. The proposed method could be applied to different interventional MRI sequences and imaging parameters by adjusting the simulation steps or by including the desired MRI parameters as inputs to the framework. Secondly, *in vivo* interventional MRI datasets were not included to evaluate the proposed technique. Due to difficulties of ground truth labeling of the physical needle position on *in vivo* datasets for training and testing, unsupervised or weakly supervised training strategies may be needed. Future work can acquire *in vivo* interventional MRI datasets from well-controlled animal experiments to further improve and evaluate our proposed technique. Thirdly, although the 3-slice model significantly reduced the through-plane physical needle localization error, the overall accuracy of the 3D physical needle localization was not yet at the level of the sub-pixel 2D localization results from the single-slice model for perfectly aligned imaging planes. Therefore, improving the 3-slice model may require a weighting factor to emphasize features on a specific channel location that is more representative of spatial information in the through-plane direction.

V. CONCLUSION

In summary, we have developed a new physical needle localization framework based on physics-driven Mask R-CNN for MRI-guided percutaneous interventions. No previous works have used the passive needle feature on MRI in a deep learning method to localize the underlying physical needle position. By calibrating the needle susceptibility value, the physics-based simulated needle feature achieved close agreement with actual MRI scans of the physical needle. We trained a single-slice physical needle Mask R-CNN model, in which the imaging plane is perfectly aligned with the needle. The testing results in *ex vivo* tissue phantoms demonstrated sub-millimeter accuracy of physical needle localization with real-time processing. In addition, we trained a 3-slice physical needle Mask R-CNN model, in which the

imaging plane is not aligned with the needle. The testing results in *ex vivo* tissue phantoms demonstrated improved through-plane physical needle localization accuracy compared with the single-slice model. Overall, the proposed framework can help to overcome the discrepancy between the passive needle feature and the physical needle during interventional MRI procedures.

ACKNOWLEDGMENT

REFERENCES

[1] S. Gupta and D. C. Madoff, "Image-guided percutaneous needle biopsy in cancer diagnosis and staging," *Techn. Vascular Interventional Radiol.*, vol. 10, no. 2, pp. 88–101, Jun. 2007, doi: 10.1053/j.tvir.2007.09.005.

[2] J. McWilliams, E. Lee, S. Yamamoto, C. Loh, and S. Kee, "Image-guided tumor ablation: Emerging technologies and future directions," *Seminars Interventional Radiol.*, vol. 27, no. 3, pp. 302–313, Sep. 2010, doi: 10.1055/s-0030-1261789.

[3] A. E. Campbell-Washburn, A. Z. Faranesh, R. J. Lederman, and M. S. Hansen, "Magnetic resonance sequences and rapid acquisition for MR-guided interventions," *Magn. Reson. Imag. Clinics North Amer.*, vol. 23, no. 4, pp. 669–679, Nov. 2015, doi: 10.1016/j.mric.2015.05.006.

[4] A. E. Campbell-Washburn, M. A. Tavallaei, M. Pop, E. K. Grant, H. Chubb, K. Rhode, and G. A. Wright, "Real-time MRI guidance of cardiac interventions," *J. Magn. Reson. Imag.*, vol. 46, no. 4, pp. 935–950, Oct. 2017, doi: 10.1002/jmri.25749.

[5] E. A. Kaye, K. L. Granlund, E. A. Morris, M. Maybody, and S. B. Solomon, "Closed-bore interventional MRI: Percutaneous biopsies and ablations," *Amer. J. Roentgenol.*, vol. 205, no. 4, pp. W400–W410, Oct. 2015, doi: 10.2214/AJR.15.14732.

[6] M. Moche, S. Heinig, N. Garnov, J. Fuchs, T.-O. Petersen, D. Seider, P. Brandmaier, T. Kahn, and H. Busse, "Navigated MRI-guided liver biopsies in a closed-bore scanner: Experience in 52 patients," *Eur. Radiol.*, vol. 26, no. 8, pp. 2462–2470, Aug. 2016, doi: 10.1007/s00330-015-4097-1.

[7] Y.-H. Lee, X. Li, J. Simonelli, D. Lu, H. H. Wu, and T.-C. Tsao, "Adaptive tracking control of one-dimensional respiration induced moving targets by real-time magnetic resonance imaging feedback," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 4, pp. 1894–1903, 2020, doi: 10.1109/TMECH.2020.2998150.

[8] M. E. Ladd, P. Erhart, J. F. Debatin, B. J. Romanowski, P. Boesiger, and G. C. McKinnon, "Biopsy needle susceptibility artifacts," *Magn. Reson. Med.*, vol. 36, no. 4, pp. 646–651, Oct. 1996, doi: 10.1002/mrm.1910360423.

[9] F. Zijlstra, J. G. Bouwman, I. Braškutė, M. A. Viergever, and P. R. Seevinck, "Fast Fourier-based simulation of off-resonance artifacts in steady-state gradient echo MRI applied to metal object localization," *Magn. Reson. Med.*, vol. 78, no. 5, pp. 2035–2041, Nov. 2017, doi: 10.1002/mrm.26556.

[10] X. Li, A. S. Young, S. S. Raman, D. S. Lu, Y.-H. Lee, T.-C. Tsao, and H. H. Wu, "Automatic needle tracking using mask R-CNN for MRI-guided percutaneous interventions," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 15, no. 10, pp. 1673–1684, Oct. 2020, doi: 10.1007/s11548-020-02226-8.

[11] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.

[12] A. Mehrtash, M. Ghafoorian, G. Pernelle, A. Ziaei, F. G. Heslinga, K. Tuncali, A. Fedorov, R. Kikinis, C. M. Tempany, W. M. Wells, P. Abolmaesumi, and T. Kapur, "Automatic needle segmentation and localization in MRI with 3-D convolutional neural networks: Application to MRI-targeted prostate biopsy," *IEEE Trans. Med. Imag.*, vol. 38, no. 4, pp. 1026–1036, Apr. 2019, doi: 10.1109/TMI.2018.2876796.

[13] J. Weine, E. Breton, J. Garnon, A. Gangi, and F. Maier, "Deep learning based needle localization on real-time MR images of patients acquired during MR-guided percutaneous interventions," in *Proc. 27th Annu. Meeting (ISMRM)*, 2019, p. 973.

[14] E. Beld, M. A. Moerland, J. R. N. van der Voort van Zyp, M. A. Viergever, J. J. W. Lagendijk, and P. R. Seevinck, "MRI artifact simulation for clinically relevant MRI sequences for guidance of prostate HDR brachytherapy," *Phys. Med. Biol.*, vol. 64, no. 9, Apr. 2019, Art. no. 095006, doi: 10.1088/1361-6560/ab15ed.

[15] S. P. DiMaio, "Needle artifact localization in 3T MR images," in *Studies in Health Technology and Informatics*, vol. 119. Amsterdam, The Netherlands: IOS Press, 2006, pp. 120–125.

[16] H. Liu, A. J. Martin, and C. L. Truwit, "Interventional MRI at high-field (1.5 T): Needle artifacts," *J. Magn. Reson. Imag.*, vol. 8, no. 1, pp. 9–214, Jan./Feb. 1998, doi: 10.1002/jmri.1880080136.

[17] S.-E. Song, N. B. Cho, I. I. Iordachita, P. Guion, G. Fichtinger, A. Kaushal, K. Camphausen, and L. L. Whitcomb, "Biopsy needle artifact localization in MRI-guided robotic transrectal prostate intervention," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 7, pp. 1902–1911, Jul. 2012, doi: 10.1109/TBME.2012.2192118.

[18] Y. K. Kim, Y. K. Kim, H. J. Park, M. J. Park, W. J. Lee, and D. Choi, "Noncontrast MRI with diffusion-weighted imaging as the sole imaging modality for detecting liver malignancy in patients with high risk for hepatocellular carcinoma," *Magn. Reson. Imag.*, vol. 32, no. 6, pp. 610–618, Jul. 2014, doi: 10.1016/j.mri.2013.12.021.

[19] L. Sonnow, W. D. Gilson, E. Raithel, M. Nittka, F. Wacker, and J. Fritz, "Instrument visualization using conventional and compressed sensing SEMAC for interventional MRI at 3T," *J. Magn. Reson. Imag.*, vol. 47, no. 5, pp. 1306–1315, May 2018, doi: 10.1002/jmri.25858.

[20] B. Müller-Bierl, H. Graf, U. Lauer, G. Steidle, and F. Schick, "Numerical modeling of needle tip artifacts in MR gradient echo imaging," *Med. Phys.*, vol. 31, no. 3, pp. 579–587, Feb. 2004, doi: 10.1118/1.1640971.

[21] T. Liu, P. Spincemaille, L. de Rochefort, B. Kressler, and Y. Wang, "Calculation of susceptibility through multiple orientation sampling (COSMOS): A method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI," *Magn. Reson. Med.*, vol. 61, no. 1, pp. 196–204, Jan. 2009, doi: 10.1002/mrm.21828.

[22] S. Bollmann, K. G. B. Rasmussen, M. Kristensen, R. G. Blendal, L. R. Østergaard, M. Plocharski, K. O'Brien, C. Langkammer, A. Janke, and M. Barth, "DeepQSM–using deep learning to solve the dipole inversion for quantitative susceptibility mapping," *NeuroImage*, vol. 195, pp. 373–383, Jul. 2019, doi: 10.1016/j.neuroimage.2019.03.060.

[23] J. Yoon, E. Gong, I. Chatnuntawech, B. Bilgic, J. Lee, W. Jung, J. Ko, H. Jung, K. Setsompop, G. Zaharchuk, E. Y. Kim, J. Pauly, and J. Lee, "Quantitative susceptibility mapping using deep neural network: QSMnet," *NeuroImage*, vol. 179, pp. 199–206, Oct. 2018, doi: 10.1016/j.neuroimage.2018.06.030.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[25] J. Stattaus, S. Maderwald, H. A. Baba, G. Gerken, J. Barkhausen, M. Forsting, and M. E. Ladd, "MR-guided liver biopsy within a short, wide-bore 1.5 Tesla MR system," *Eur. Radiol.*, vol. 18, no. 12, pp. 2865–2873, Dec. 2008, doi: 10.1007/s00330-008-1088-5.

[26] J. Stattaus, S. Maderwald, M. Forsting, J. Barkhausen, and M. E. Ladd, "MR-guided core biopsy with MR fluoroscopy using a short, wide-bore 1.5-Tesla scanner: Feasibility and initial results," *J. Magn. Reson. Imag.*, vol. 27, no. 5, pp. 1181–1187, May 2008, doi: 10.1002/jmri.21075.

[27] N. Tan, W.-C. Lin, P. Khoshnoodi, N. H. Asvadi, J. Yoshida, D. J. A. Margolis, D. S. K. Lu, H. Wu, K. H. Sung, D. Y. Lu, J. Huang, and S. S. Raman, "In-bore 3-T MR-guided transrectal targeted prostate biopsy: Prostate imaging reporting and data system version 2–based diagnostic performance for detection of prostate cancer," *Radiology*, vol. 283, no. 1, pp. 130–139, Apr. 2017, doi: 10.1148/radiol.2016152827.

[28] E. Rothgang, W. D. Gilson, W. Strehl, L. Pan, J. Roland, C. H. Lorenz, and J. Hornegger, "Interventional MR-imaging for thermal ablation therapy," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Mar. 2011, pp. 1864–1868, doi: 10.1109/ISBI.2011.5872771.

[29] X. Li, Y.-H. Lee, T.-C. Tsao, and H. H. Wu, "Physical needle localization using mask R-CNN for MRI-guided percutaneous interventions," in *Proc. 28th Annu. Meeting (ISMRM)*, 2020, p. 4144.

[30] X. Li, "Computer-assisted navigation techniques for MRI-guided percutaneous interventions," Ph.D. dissertation, Dept. Bioeng., UCLA, Los Angeles, CA, USA, 2021.

[31] W. Abdulla. (2017). *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. GitHub Repository. [Online]. Available: https://github.com/matterport/Mask_RCNN

[32] X. Li, Y.-H. Lee, S. Mikaiel, J. Simonelli, T.-C. Tsao, and H. H. Wu, "Respiratory motion prediction using fusion-based multi-rate Kalman filtering and real-time golden-angle radial MRI," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 6, pp. 1727–1738, Jun. 2020, doi: 10.1109/TBME.2019.2944803.

[33] T. S. Sorensen, D. Atkinson, T. Schaeffter, and M. S. Hansen, "Real-time reconstruction of sensitivity encoded radial magnetic resonance imaging using a graphics processing unit," *IEEE Trans. Med. Imag.*, vol. 28, no. 12, pp. 1974–1985, Dec. 2009, doi: 10.1109/TMI.2009.2027118.

[34] J. Simonelli, Y.-H. Lee, C.-W. Chen, X. Li, S. Mikaiel, D. Lu, H. H. Wu, and T.-C. Tsao, "Hydrostatic actuation for remote operations in MR environment," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 2, pp. 894–905, Apr. 2020, doi: 10.1109/TMECH.2019.2959805.

[35] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, and J. Buatti, "3D slicer as an image computing platform for the quantitative imaging network," *Magn. Reson. Med.*, vol. 30, no. 9, pp. 41–1323, Nov. 2012, doi: 10.1016/j.mri.2012.05.001.

[36] R. Salomir, B. D. de Senneville, and C. T. W. Moonen, "A fast calculation method for magnetic field inhomogeneity due to an arbitrary distribution of bulk susceptibility," *Concepts Magn. Reson. B, Magn. Reson. Eng.*, vol. 19B, no. 1, pp. 26–34, Jan. 2003, doi: 10.1002/cmr.b.10083.

[37] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 560–574, Feb. 2003, doi: 10.1109/TSP.2002.807005.

[38] Y. L. Park, S. Elayaperumal, B. Daniel, S. C. Ryu, M. Shin, J. Savall, R. J. Black, B. Moslehi, and M. R. Cutkosky, "Real-time estimation of 3-D needle shape and deflection for MRI-guided interventions," *IEEE/ASME Trans. Mechatronics*, vol. 15, no. 6, pp. 906–915, Dec. 2010, doi: 10.1109/TMECH.2010.2080360.

[39] E. Franco, M. Ristic, M. Rea, and W. M. W. Gedroyc, "Robot-assistant for MRI-guided liver ablation: A pilot study," *Med. Phys.*, vol. 43, no. 10, pp. 5347–5356, Sep. 2016, doi: 10.1118/1.4961986.

[40] N. Hata, S.-E. Song, O. Olubiyi, Y. Arimitsu, K. Fujimoto, T. Kato, K. Tuncali, S. Tani, and J. Tokuda, "Body-mounted robotic instrument guide for image-guided cryotherapy of renal cancer," *Med. Phys.*, vol. 43, no. 2, pp. 843–853, Jan. 2016, doi: 10.1118/1.4939875.

[41] X. Li, Y.-H. Lee, D. Lu, T.-C. Tsao, and H. H. Wu, "Deep learning-driven automatic scan plane alignment for needle tracking in MRI-guidedInterventions," in *Proc. 29th Annu. Meeting (ISMRM)*, 2021, p. 861.

YU-HSIU LEE received the B.S. and M.S. degrees in mechanical engineering from the National Taiwan University, Taipei, Taiwan, in 2007 and 2009, respectively, and the Ph.D. degree in mechanical and aerospace engineering from the University of California Los Angeles, Los Angeles, USA, in 2019.

He is currently an Assistant Professor of mechanical engineering at the National Taiwan University.

DAVID S. LU received the M.D. degree from McGill University, Montreal, QC, Canada, in 1985.

He was a Radiology Resident with the University of Toronto, in 1991, and a Radiology Fellow with Massachusetts General Hospital, Harvard University, in 1993. He is currently the Director of interventional oncology and cross sectional interventional radiology at the University of California Los Angeles, Los Angeles, CA, USA. His research interests include ablation, interventional MRI, navigation tools, such as fusion imaging through electromagnetic tracking, and protective techniques to enhance ablation safety.

TSU-CHIN TSAO (Senior Member, IEEE) received the B.S. degree in engineering from the National Taiwan University, Taipei, Taiwan, in 1981, and the M.S. and Ph.D. degrees in mechanical engineering from the University of California at Berkeley, Berkeley, CA, USA, in 1984 and 1988, respectively.

He is currently a Professor with the Mechanical and Aerospace Engineering Department, University of California Los Angeles, Los Angeles, CA, USA. His research interests include precision motion control, mechatronics, and robotics.

XINZHOU LI received the B.S. degree in bioengineering and mechanical engineering from Pennsylvania State University, State College, PA, USA, in 2014, and the Ph.D. degree in bioengineering from the University of California Los Angeles (UCLA), Los Angeles, USA, in 2021.

He is currently a Postdoctoral Researcher with UCLA Radiological Sciences, where the focus of his study is real-time MRI and MRI-guided interventions.

HOLDEN H. WU (Member, IEEE) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 2003, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 2005 and 2009, respectively.

He is currently an Associate Professor of radiological sciences and bioengineering with the University of California Los Angeles, Los Angeles, CA, USA. His research interests include the development and translation of quantitative magnetic resonance imaging (MRI) and MRI-guided interventions for cancer and metabolic diseases.

• • •