Rupture Detection During Needle Insertion Using Complex OCT Data and CNNs

Sarah Latus, Johanna Sprenger, Maximilian Neidhardt, Julia Schädler, Alexandra Ron, Antonia Fitzek, Matthias Schlüter, Philipp Breitfeld, Axel Heinemann, Klaus Püschel, and Alexander Schlaefer

Abstract*—*Objective: **Soft tissue deformation and ruptures complicate needle placement. However, ruptures at tissue interfaces also contain information which helps physicians to navigate through different layers. This navigation task can be challenging, whenever ultrasound (US) image guidance is hard to align and externally sensed forces are superimposed by friction.** ***Methods:*** **We propose an experimental setup for reproducible needle insertions, applying optical coherence tomography (OCT) directly at the needle tip as well as external US and force measurements. Processing the complex OCT data is challenging as the penetration depth is limited and the data can be difficult to interpret. Using a machine learning approach, we show that ruptures can be detected in the complex OCT data without additional external guidance or measurements after training with multi-modal ground-truth from US and force.** ***Results:*** **We can detect ruptures with accuracies of 0.94 and 0.91 on homogeneous and inhomogeneous phantoms, respectively, and 0.71 for ex-situ tissues.** ***Conclusion:*** **We propose an experimental setup and deep learning based rupture detection for the complex OCT data in front of the needle tip, even in deeper tissue structures without the need for US or force sensor guiding.** ***Significance:*** **This study promises a suitable approach to complement a robust robotic needle placement.**

***Index Terms*—Deep learning, needle navigation, optical coherence tomography, relative tissue motion.**

I. Introduction

NEEDLE placement is a challenging task in different medical applications. For example during biopsies, precise localization of needle and tumor boundaries as well as reliable placement of the needle tip within the tumorous tissue are important when sampling the tissue. Another example is the placement of needles to administer epidural anesthesia, where the physician penetrates a number of tissue structures and needs to stop directly behind the ligamentum flavum to reach the small epidural space of a few millimeters length. An accidental puncture of the dura might cause a critical liquor loss [1], [2]. Generally, several tissue structures with varying elasticities are punctured during needle insertions. Experienced physicians typically rely on the haptic impression at the needle shaft to navigate. Sometimes they can identify tissue structures, as ruptures of tissue layers can be felt at the shaft, e.g., during liver biopsies [3].

However, forces at the tip and the haptic impression caused be a sequence of ruptures during needle insertion become harder to interpret for deeper insertions, where friction acting on the shaft increases [4], [5]. Hence, it becomes virtually impossible to sense the interaction of the needle with the tissue structures at its tip at the distal shaft. Moreover, the tissue is deformed and displaced by the needle [6], making it difficult to identify the exact position of the needle tip relative to the tissue and the target by only observing the shaft motion. Even if the needle is placed by a robot [7], the position of the needle relative to soft tissue targets is often difficult to assess. Hence, accurate detection of tissue boundaries and ruptures during needle insertions is required to allow for accurate navigation robust to relative motion.

Approaches to estimate the needle pose include optical [8], [9] and electromagnetic tracking [10]. Furthermore, image guidance by e.g. computed tomography, magnetic resonance imaging and especially ultrasound (US) imaging have been studied. During 2D US imaging, out-of-plane motion by the needle limits the needle pose estimation. A more stable needle tracking can be achieved with echogenic needles [11], coded excitation [12], or 3D US probes [13]. However, US cannot be used for deeper tissue structures and those shadowed by bones.

Due to the lack of sufficient image guidance and to overcome the effect of superimposed friction forces during insertion, additional sensors or miniaturized imaging probes can be embedded in needles to measure tissue properties directly at the needle tip. For example, an estimation of the tip forces can be achieved by fiber Bragg gratings [14] or Fabry-Pérot interferometry [15]. Alternatively, optical fibers can be embedded to enable optical reflectance spectroscopy [16], Raman spectroscopy [17], or OCT imaging of the interacting tissue directly at the needle tip. OCT provides a high spatial and temporal resolution for analyzing small tissue structures [18]. The

Manuscript received September 18, 2020; revised February 2, 2021; accepted February 19, 2021. Date of publication March 2, 2021; date of current version September 20, 2021. This work was supported by the i^3 initiative of the Hamburg University of Technology, and by the German Research Foundation DFG, under Grant SCHL 1844/2-1/2. *(Sarah Latus and Johanna Sprenger contributed equally to this work.) (Corresponding author: Sarah Latus.)*

Sarah Latus is with the Hamburg University of Technology, Institute of Medical Technology and Intelligent Systems, 21071 Hamburg, Germany (e-mail: s.latus@tuhh.de).

Johanna Sprenger, Maximilian Neidhardt, Matthias Schlüter, and Alexander Schlaefer are with the Hamburg University of Technology, Institute of Medical Technology and Intelligent Systems, Germany.

Julia Schädler, Alexandra Ron, Antonia Fitzek, Axel Heinemann, and Klaus Püschel are with the University Medical Center Hamburg-Eppendorf, Institute of Forensic Medicine, Germany.

Philipp Breitfeld is with the University Medical Center Hamburg-Eppendorf, Department of Anesthesiology, Germany.

Digital Object Identifier 10.1109/TBME.2021.3063069

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

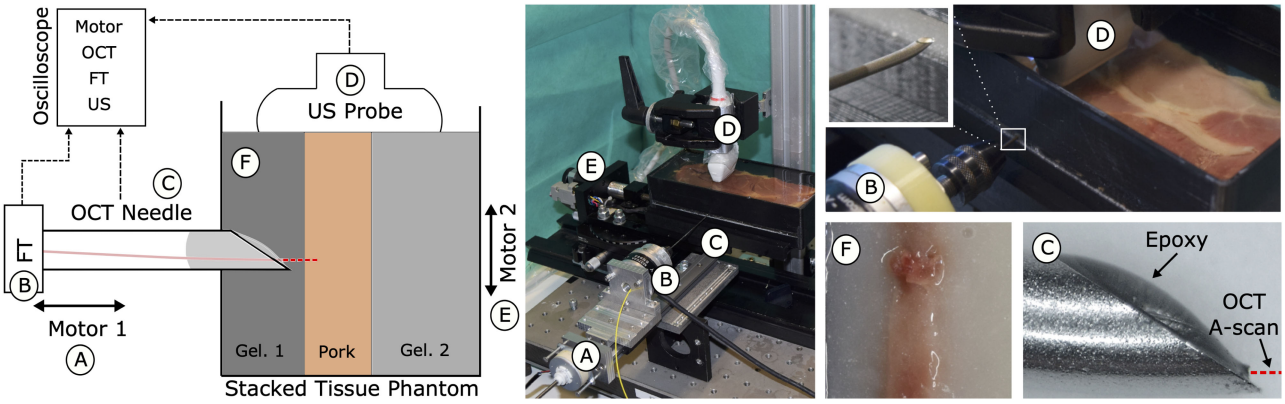

Fig. 1. Experimental setup for phantom studies. The forward-viewing OCT needle (C) is inserted into a stacked tissue phantom (F) using a linear motion stage (A, motor 1). A linear US probe (D) is positioned perpendicular to the needle trajectory. The FT sensor (B) is mounted on the needle shaft. The trigger signals of all systems are collected synchronously at the oscilloscope. A second motion stage (E, motor 2) performs a phantom reorientation after puncture.

morphological information from OCT intensity data can be used to characterize [19] and classify [20] different tissue structures. Additionally, needle based OCT elastography allows to estimate mechanical tissue properties using OCT phase data [21], [22].

Force sensing at the needle tip [23] has been facilitated with a combination of OCT and deep learning methods. Nevertheless, this approach only allows to analyze the motion of an epoxy resin layer attached to the needle tip, but tissue imaging in front of the needle has not been considered. Applications using the different OCT needle designs [19]-[23] require a fast processing and classification of the OCT image data for online control which is challenging due to the high amount of data and the tissue dependent signal characteristics.

Deep learning has also been proposed to identify tissue structures in OCT data. Previous applications include segmentation of the retina [24], automatic plaque detection in arteries [25], or classification and analyzation of ophthalmic structures [26].

We study OCT imaging at the needle tip and deep learning methods to detect rupture events during needle insertions. Particularly, we consider sensing the needle tissue interactions and the ruptures caused by inserting the needle at its tip by feeding the OCT signal into a convolutional neural network (CNNs). While CNNs have become a standard approach for image processing, we adapt our network to consider the complex OCT signal, i.e., magnitude and phase. To obtain labeled data we present an experimental setup employing ultrasound and force sensors as independent surrogates. The needles were moved by a linear stage to obtain reproducible measurements. Data was acquired for different tissue samples with varying elastic properties and we provide systematic evaluation with respect to the different CNN input and the surrogate signals. Our results demonstrate that ruptures of tissue at the needle tip can be detected from within the embedded fiber, providing additional information for navigation when image guidance is not feasible, e.g., in deep tissue structures.

II. Material and Methods

In this chapter, we first introduce our experimental setup and preprocessing methods applied to our multi-modal data sets. Second, our deep learning methods for OCT-based rupture detection are proposed. Third, the experimental studies, punctured tissue phantoms and ex-situ tissue samples are described.

A. Experimental Setups

1) Phantom Setup: Our phantom setup for reproducible needle insertions in tissue phantoms is composed of a forward-viewing OCT needle, a needle motion stage, a force-torque (FT) sensor, a linear US probe, a phantom motion stage, and an oscilloscope as depicted in Fig. 1. Our OCT system (Telesto I, Thorlabs) allows to acquire one-dimensional depth-resolved scans (A-scans) with a scan rate of 91kHz and imaging depth of 512 pixels which is equivalent to 2.56mm in air. In tissue structures with a refractive index of approximately 1.45, this imaging depth results in about 1.77mm. Our forward-viewing OCT needle design is shown in Fig. 1 C. The optical fiber tip is cleaved to obtain a perfectly flat endface and facilitates a common-path A-scan acquisition. The fiber is positioned at the tip of an epidural Tuohy needle with 1.4mm diameter and glued with epoxy adhesive resin. Using this design, the OCT A-scan is oriented forward-viewing along the needle axis (Fig. 1 C). We use a FT sensor (M3703 A, Sunrise Instruments) to acquire force measurements with a frequency of up to 2kHz. We connect a linear US probe (CPLA128, Cephasonics Ultrasound) with a center frequency of 10MHz to an US system (Griffin, Cephasonics Ultrasound) and receive two dimensional images (128×364 px) with a frame rate of 65fps. Assuming a homogeneous speed of sound of 1540 mms^{-1}, we obtain a field-of-view (FOV) of approximately 27×40 mm.

The needle insertion and multi-modal data acquisition are controlled as follows: We attach the OCT needle to a motion stage and continuously acquire OCT A-scans during the needle insertion. In addition, external forces at the needle shaft are measured. The US probe is attached rigidly to the needle motion stage and enables an in-plane tracking of the needle motion and tissue deformation in front of the needle tip. The needle is mechanically guided to limit needle deflections and enable a needle motion parallel to the lateral axis of the US FOV. The oscilloscope (Logic Pro 8, Saleae) is applied to synchronously

collect the trigger signals of all components (US, FT, OCT, motor steps) with a sampling rate of 325kHz. After each insertion, a second motion stage rearranges the phantom.

Due to the limited lateral FOV (27mm) of our US imaging probe we need to crop our US, OCT, and FT data sets over time regarding to the entry and exit of the needle in the US frame. Hence, we calibrate our phantom setup and determine the needle trajectory within the 2D+t US data for 40 needle punctures in transparent phantoms of homogeneous density. The needle position is determined in each US frame using a combination of a threshold based needle segmentation and needle velocity estimation derived from the Lucas-Kanade method [27].

2) Ex-Situ Setup: In our ex-situ setup (Fig. 5) the OCT system, forward-viewing OCT needle, FT sensor, needle motion stage as well as the oscilloscope are controlled similarly. An US probe is not mounted due to the limited space in this setup. During needle insertion we simultaneously acquire OCT A-scans, force values measured at the needle shaft and log the trigger signals from OCT, FT sensor and motion stage at the oscilloscope.

B. Rupture Detection

In the following, we describe the processing of the acquired multi-modal data to extract tissue displacement profiles. Considering the tissue displacements over time, we define a rupture event as a point in time, where the needle punctures a tissue boundary and enters the underlying tissue structure.

1) OCT Image Processing: The complex OCT data of each needle insertion carries information about the compression and strain of tissue in front of the needle tip and is separated into magnitude (intensity) and phase data. In Fig. 2 an exemplary OCT M-scan of a pork-gelatin puncture is shown. The A-scans are displayed side by side over time. Note, these M-scans can not be interpreted as 2D cross-sectional images of the punctured tissue. Tissue structures are compressed with different velocities towards the needle tip and tissue boundaries appear as lines with varying slopes in the OCT M-scan. We exemplarily highlighted structures that are visible in sequential A-scans with red and blue. The gradients of the structures are related to the tissue compression and puncture during needle insertion.

To derive the tissue displacement, we estimate the relative tissue speed v [22] in front of the needle based on the OCT phase data. We apply phase unwrapping and use the phase difference $\Delta\varphi$ of successive A-scans to calculate

$$v = \frac{\Delta\varphi \cdot \lambda_0}{4\pi T}. \tag{1}$$

The time difference between A-scans T is measured at the oscilloscope and the center wavelength $\lambda_0 = 1300nm$ is given from the OCT system specifications. We extract rupture events t_{oct} as peaks with $v > p_{\text{oct}} v_{\text{m}}$, where v_{m} is defined as maximum peak height per insertion and p_{oct} as scaling factor. We determine the prominence for all local maxima in the displacement profiles, whereas the greatest prominence is defined as maximum peak height per insertion.

2) US Image Processing: Fig. 3 shows the US data of an exemplary puncture of a pork tissue-gelatin boundary. The tissue

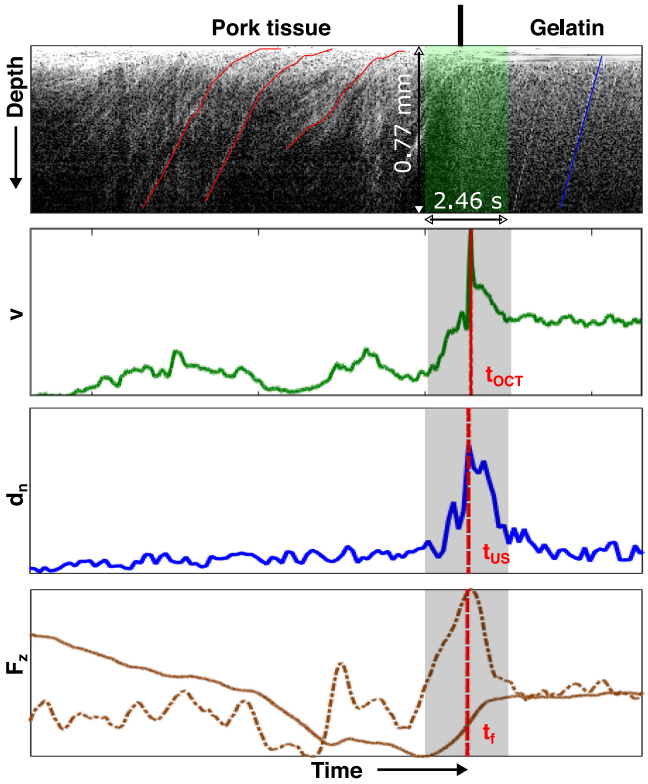

Fig. 2. Multi-modal data from a pork-gelatin puncture. OCT intensity data with different OCT speckle characteristics from pork tissue and gelatin is shown on top. The passing structures of the pork tissue and TiO2 scatters in the gelatin are highlighted exemplary in red and blue, respectively. The displacement profiles from OCT phase data (v, green), US data (d_{n}, blue), force (F_z, brown) and related gradient (∇F_z, dashed brown) are plotted over time. The extracted rupture events (t_{oct}, t_{us}, and t_{f}) are highlighted with red dashed lines. The deep learning intensity input around the rupture of 224×224 pixels which is equivalent to 2.46s is highlighted with a green box, the related time intervals in the other modalities in gray.

boundary deforms due to the moving needle tip in positive x-direction. We extract deformation profiles from US 2D+t data by estimating the optical flow of particles in succeeding US frames based on the Lucas-Kanade method [27]. We quantify the boundary deformation d by calculating the mean magnitude of the optical flow vectors $\vec{o}_{i,j}$ in a subset ($N \times M$ px) of the US frame

$$d = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} |\vec{o}_{i,j}|. \tag{2}$$

To consider the tissue deformation direction with respect to the needle motion, we differentiate between positive and negative deformation profiles related to the optical flow vector orientations in positive (red) and negative (blue) x-direction (Fig. 3, bottom). We extract a rupture event t_{us} from the negative deformation profile as a peak with height of $|d_{\text{n}}| > p_{\text{us}}|d_{\text{m}}|$, with d_{m} as maximum peak value per insertion and the scaling factor p_{us}. In Fig. 3 this peak can be related to the turned optical flow vectors at position G.

3) Force Data Processing: We assume a constant friction at the needle shaft, while the force in needle direction F_z

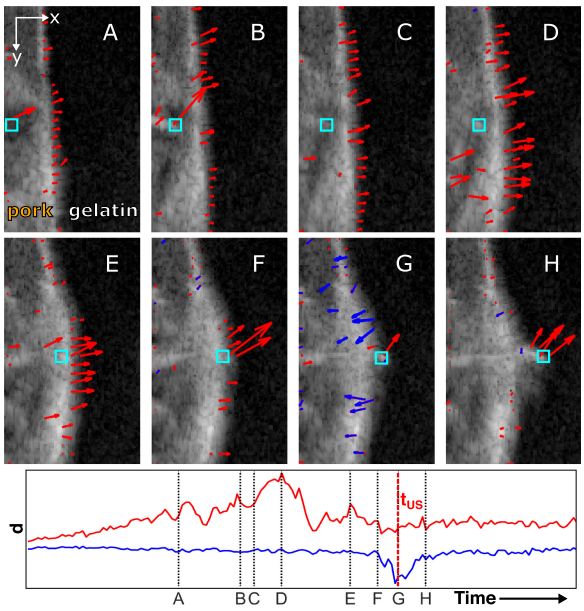

Fig. 3. (Top) Exemplary needle puncture of a pork tissue and gelatin boundary. In the cropped US images (10×15 mm^2) over time, the related optical flow vectors are depicted in red for different deformation states from needle entrance (A) to the boundary rupture (G) and following needle insertion in gelatin (H). The needle tip is boxed in cyan. (Bottom) Related deformation profiles (d) over time with depicted time points of shown US frames (A-H). The deformation profiles in positive and negative x-direction are plotted in red and blue, respectively. The rupture t_{us} at the tissue-gelatin boundary can be related to the negative peak in the blue curve (G).

increases linearly with insertion depth. Within our experimental setup, we assume that an abruptly inverting slope of F_z is related to a rupture at a tissue boundary. To extract tissue displacement profiles, we smooth the force data by applying a moving average and calculate the gradient ∇F_z of the detected force values in needle direction. We extract rupture events t_{f} as prominent peaks with a peak height of $\nabla F_z > p_{\text{f}} \nabla F_{z,m}$. Here, the maximum peak height $\nabla F_{z,m}$ per insertion is used as reference and p_{f} is denoted as scaling factor.

For each modality, we analyze different scaling factors p to extract rupture events from the tissue displacement profiles. In the following, this is mentioned as threshold based rupture extraction method. An exemplary rupture labeling for a tissue-gelatin boundary puncture is shown in Fig. 2. The displacement profiles are synchronized using the trigger signals captured with the oscilloscope.

C. Deep Learning Methods

1) Architectures: We consider a pretrained ResNet-50 [28] model for our approach to detect ruptures in OCT data. As the data set is comparatively small and previous work has shown the effectiveness of transfer learning in the medical imaging domain [29], we employ a model pretrained on ImageNet data [30]. We enable weight updates for the last few layers to adapt the model to the OCT data. Furthermore, we replace the final fully-connected layer to match our binary classification task. The pretrained models require three-channel images of size 224×224 pixels. The intensity and unwrapped phase data samples of the complex OCT M-scans are merged to three-channel images each and passed to the models. We evaluate different input data sets composed of intensity or phase data.

For combining intensity and phase of the complex OCT data, we introduce a method based on using a support vector machine (SVM) [31]. This combined approach is visualized in Fig. 4. We train two separated models on intensity and phase data, respectively. The SVM is used for a final rupture prediction n_{r} based on the models outputs n_i. Therefore, we pass the training, validation and test data through the networks to generate the probability estimations. The training and validation outputs of the two models are used as input features to fit a SVM. The SVM extracts the essential information to predict the output class. Subsequently, we pass the CNNs test data probability estimations, based on intensity and phase data, to the SVM to classify the samples.

2) Input Data and Labels: The OCT A-scans are cropped to obtain a depth of 224 pixels in the axial direction. We keep the pixels closest to the needle and cut pixels in A-scan depth. We create the CNN input composed of 2D-image subsets of the OCT M-scans, whereby the M-scans are previously reduced to every 1000th A-scan. Considering the different sampling rates of our measurement modalities, a rupture event might be detected at slightly different points in time, as shown in Fig. 2. This is due to the different sampling rates of the ground-truth modalities. The force and A-scan acquisition rates are sub-sampled to match the lower US imaging frequency. This avoids the need of interpolating the signals and inducing inaccuracies. We use the determined rupture events as labels and extract images of size 224×224 pixels around the label's position to match the standard architecture's input sizes [32]. Hence, we do not use a certain point of time to localize the rupture, but rather detect the occurrence of a rupture in the given time interval. The data is sampled around the determined rupture event and therefore slight differences in the in the multi-modal measurements to generate the ground-truth do not effect the ground-truth labels itself. Our binary labels therefore encode the states rupture and no rupture either using the US ground-truth t_{us} or force ground-truth t_{f}. Furthermore, we use the multi-modal labels t_{us} *and* t_{f} (AND) and t_{us} *or* t_{f} (OR) to evaluate the influence of the two ground-truth modalities. The former is positive only if a rupture is labeled both in US and force. The latter is positive if either, US or force, is labeled with a rupture. During training, the networks are presented the four labels to predict ruptures based on the complex OCT data.

D. Experiments

We perform several insertion experiments to evaluate the detection of ruptures in the complex OCT data based on standard image processing (Section II-B1) and deep learning (Section II-C) methods. In the following, we introduce the preparation of differently composed tissue phantoms, ex-situ tissue samples [33] and the applied experimental parameters.

1) Phantom Study: For our phantom study, we prepare homogeneous and inhomogeneous tissue phantoms. The homogeneous phantoms consist of pork filet without fat, sinews

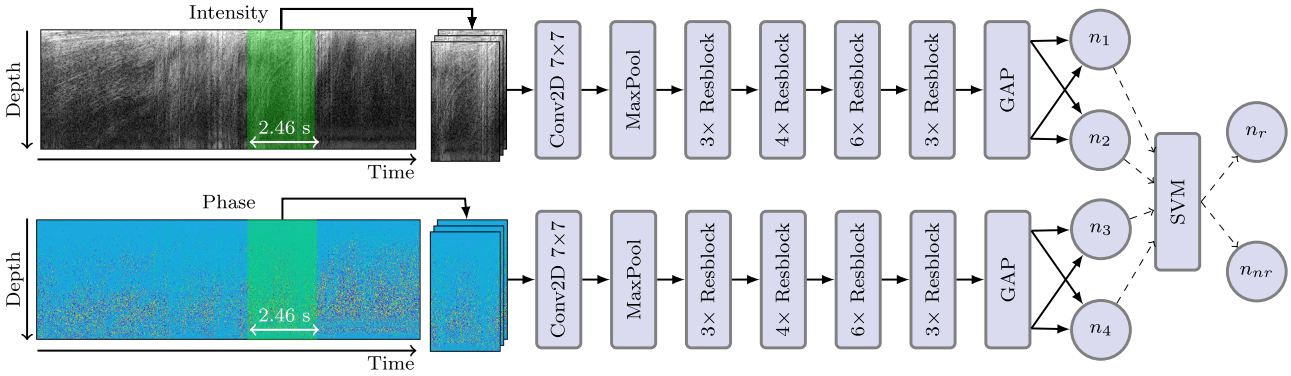

Fig. 4. Our approach for preparing the complex OCT data as input for pretrained models and using both intensity and phase data combined. Patches (denoted by the green overlay) are extracted from the OCT intensity and phase data at corresponding positions and passed as three-channel images to the pretrained ResNet-50 models. The models are trained on intensity and phase data, respectively, and the outputs n_i are passed to a SVM for a final combined prediction of a rupture n_r or non-rupture event n_nr.

TABLE I
PREPARED HOMOGENEOUS (HOM.) AND INHOMOGENEOUS (INHOM.) TISSUE PHANTOMS AND EX-SITU TISSUE SAMPLES. FOR THE TISSUE PHANTOMS, THE DENSITIES OF THE GELATIN IN FRONT (GEL. 1) AND BEHIND (GEL. 2) THE PORK TISSUE ARE GIVEN AS RATIO OF GELATIN POWDER TO WATER (GELATIN:WATER). FOR EACH PHANTOM AND EX-SITU TISSUE TYPE THE NUMBER OF PUNCTURES AND PREPARED SAMPLES ARE GIVEN

Name	Type	Gel. 1	Gel. 2	punctures	samples
A	hom.	1:15	1:15	33	2
B		1:15	1:5	33	2
C		1:5	1:15	33	2
D	inhom.	1:5	1:5	51	4
liver	ex-situ	–	–	5	5
prostate		–	–	4	4

or muscles, embedded in gelatin. The phantoms are composed of three layers (gelatin-tissue-gelatin, Fig. 1 F) and are punctured perpendicular to the almost straight pork tissue-gelatin boundaries. To simulate different tissue elasticities, we vary the density of the gelatin in front of (Gel. 1) and behind (Gel. 2) the pork filet by mixing different concentrations of gelatin powder to water. In total, three differently stacked homogeneous phantom types are used (Table I). First, phantoms without a density gradient of the gelatin in front of or behind the pork are prepared (type A). Second, a tumor formation is simulated by means of hard gelatin behind the pork tissue (type B). Third, soft gelatin behind the pork tissue is used to simulate a loss-of-resistance, as typically occurring during epidural punctures (type C). Particles of titanium dioxide are added to the gelatin as optical scatters for OCT imaging. The inhomogeneous tissue phantoms (type D) are composed of sinewy pork filet embedded in a hard gelatin mixture (Gel. 1 = Gel. 2 = 1:5, Fig. 1 top right). The filet is positioned horizontal to the needle trajectory to enable a puncture of sinews, muscle, and fat layers. As the gelatin is only used to fixate the sample and the image data of the gelatin pork boundaries are not processed, we do not add optical scatters in the gelatin of these phantoms.

The needle motion stage of our phantom setup is set to a constant insertion velocity of 1 mms^{-1} over a total distance of 56mm. After each insertion, the phantom motion stage rearranges the phantom to obtain a distance of 10mm between the needle trajectories. Using a phantom width of 200mm, in total up to 18 punctures per phantom are conducted. In Table I, the resulting number of punctures per phantom type and number of prepared phantom samples are listed. We perform in total 150 needle punctures in differently layered homogeneous and inhomogeneous gelatin-tissue phantoms (see Table I).

2) Ex-Situ Study: Additionally, we investigate ex-situ tissue samples gathered during an autopsy. The tissue samples are examined by forensic pathologists macroscopic to determine if pathological lesions are present. The tissues are cut to a size of $20 \times 20 \times 15$ mm^3. An enclosing fixation of the samples is necessary, while an embedding with gelatin is not possible and a tissue compression needs to be minimized. Hence, a 3D-printed case is used to fix the tissue during puncture (Fig. 5). We punctured in total five liver tissue samples with different degenerative states. Additionally, four prostate tissue samples with homogeneous and varying tissue densities and different amounts of tissue indurations are examined. A constant insertion velocity of 1 mms^{-1} over a total distance of 25mm is applied. Due to the limited tissue dimensions, only a single puncture per tissue sample is possible (Table I). After each puncture, we cut the tissue along the needle shaft (Fig. 5 C). The tissue was fixed in buffered 4% formaldehyde. The formalin-fixed paraffin embedded (FFPE) tissue samples were processed and stained with hematoxylin and eosin (H&E) following standard laboratory procedures. Slides were examined by experienced morphologists, and selected medium were electronically scanned at high magnification (20x) as high resolution images (0.46 um / px), using a digital slide scanner (NanoZoomer 2.O-HT, Hamamatsu).

For both phantom and ex-situ study, the insertion velocity, OCT A-scan rate and the sub-sampling for deep learning data processing results in a input size of 224×224 pixels which is equivalent to approximately 0.77mm $\times$ 2.46s (A-scan depth $\times$ needle motion) as highlighted green in Fig. 2.

III. RESULTS

A. Rupture Extraction Thresholds

We apply a grid search to extract an optimal scaling parameter p_oct for OCT phase based rupture detection II-B1. We evaluate the receiver operating characteristics (ROC) for different combinations of the rupture extraction scaling factors p for all

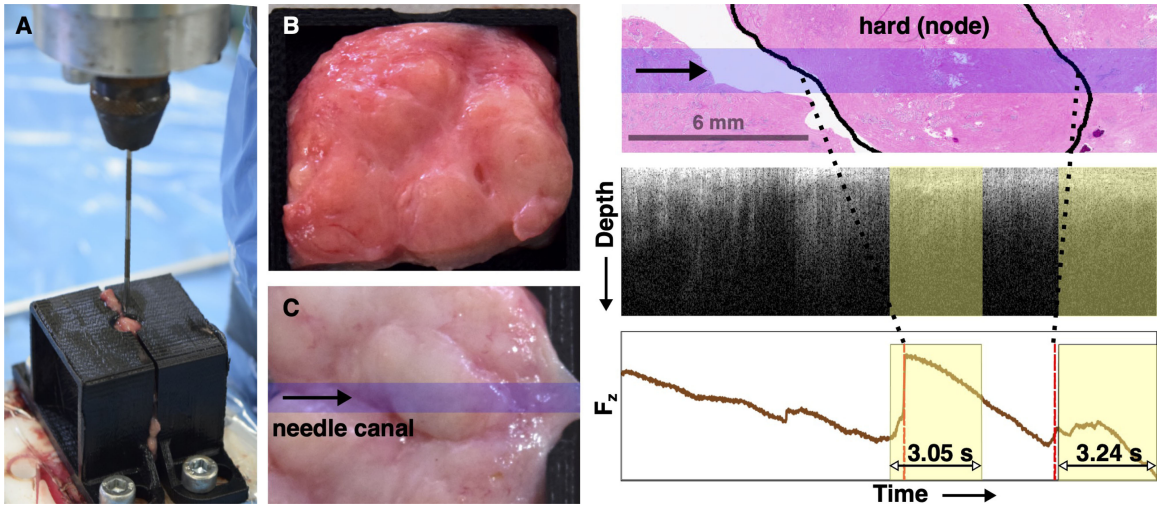

Fig. 5. (Left) Experimental setup for ex-situ studies (A). The punctured prostate tissue prior to insertion (B) and after cutting (C) is shown. (Right) Comparison of histology, OCT intensity data and force measurements F_z of an exemplary prostate tissue puncture. Tissue boundaries (black) and the needle canal (blue) are highlighted in the histology. Related rupture labels t_f are highlighted with red vertical lines. The network predictions are highlighted with yellow boxes. The histology image of the prostate tissue sample shows a hard node (stromal hyperplasia) highlighted in the center.

TABLE II
EVALUATION OF THRESHOLD BASED RUPTURE DETECTION COMPARED TO GROUND-TRUTH (GT) FROM US t_US AND FORCE t_F FOR ALL TISSUE PHANTOMS (ALL) AND SEPARATED PHANTOM TYPES (A-D). ACCURACY, PRECISION, RECALL, AND F1-SCORE ARE GIVEN

phantoms	modality	GT	Acc.	Prec.	Rec.	f1-Score
all	t_oct	t_us	0.810	0.847	0.619	0.715
	t_oct	t_f	0.828	0.770	0.680	0.718
	t_us	t_f	0.849	0.722	0.871	0.789
A	t_oct	t_us	0.940	0.988	0.862	0.921
	t_oct	t_f	0.849	1.000	0.702	0.825
	t_us	t_f	0.866	0.958	0.770	0.854
B	t_oct	t_us	0.821	0.749	0.637	0.688
	t_oct	t_f	0.793	0.497	0.639	0.559
	t_us	t_f	0.845	0.582	0.880	0.700
C	t_oct	t_us	0.730	0.776	0.499	0.607
	t_oct	t_f	0.825	0.697	0.668	0.682
	t_us	t_f	0.828	0.630	0.940	0.755
D	t_oct	t_us	0.810	0.875	0.616	0.723
	t_oct	t_f	0.838	0.828	0.678	0.745
	t_us	t_f	0.859	0.756	0.880	0.813

modalities using an interval of $p_i \in [0.2, 0.8]$. First, we determine the scaling factors for the force and US rupture extraction. Here, we choose values to obtain a maximized accordance of US and force based rupture labels. The values $p_\mathrm{f} = 0.55$ and $p_\mathrm{us} = 0.5$ follow. Second, we use these scaling factors and select the optimal scaling factor for OCT rupture extraction based on the relative speed v. A scaling factor $p_\mathrm{oct} = 0.65$ is determined from the related ROC curves. Finally, the performance of the threshold based rupture extractions are compared for different ground-truth (GT) rupture events from US and force (Table II). The accuracy (Acc.), precision (Prec.), recall (Rec.), and f1-Score are given.

B. Quantification of Rupture Events

We quantify the rupture events by means of the underlying force values. For each of the 224×224 sized inputs, we calculate the difference of the maximum and minimum force value (ΔF).

TABLE III
QUANTIFICATION OF RUPTURE EVENTS. THE MEAN AND STANDARD DEVIATIONS OF THE FORCE DIFFERENCE (ΔF, IN N) PER INPUT ARE LISTED PER TISSUE AND EVENT (RUPTURE VS. NO RUPTURE)

Name	ΔF, rupture	ΔF, no rupture
A	0.64 ± 0.58	0.13 ± 0.17
B	0.77 ± 1.28	0.21 ± 0.30
C	0.45 ± 0.42	0.18 ± 0.21
D	0.95 ± 1.01	0.43 ± 0.35
liver	0.22 ± 0.18	0.16 ± 0.16
prostate	0.87 ± 0.33	0.75 ± 0.19

We define that a small difference between the maximum and minimum force value leads to a small rupture and a larger difference to a large rupture. We distinguish between rupture and no rupture events and calculate the mean and standard deviation for each phantom type (A, B, C, D) and the ex-situ tissue samples (Table III). The phantom mixtures with increased gelatin stiffness behind the pork tissue lead to larger ruptures (phantoms B and D). An increased stiffness in front of the pork tissue decreases the rupture magnitude.

C. Deep Learning Performance

1) Phantom Study: We evaluate our network for different OCT inputs (int, phase, combined) and ground-truth labels. In Table IV, the results for CNN based rupture detection trained on the whole phantom data set (all) are shown. The first column visualizes the input and the second column contains the information about the ground-truth label (t_us, t_f, AND, OR). Based on Table IV, using phase data as input leads to better results than using intensity as input. Comparing the accuracies for both inputs over the different ground-truth labels, a difference of over 0.22 percentage points can be noticed. The improvements for phase input are also reflected in precision, recall and f1-score. Our approach for combining intensity and phase as input (combined) leads to similar scores as using phase data separately. The

TABLE IV
EVALUATION OF THE CNN BASED RUPTURE DETECTION TRAINED ON THE WHOLE DATA SET. DIFFERENT INPUTS AND GROUND-TRUTH ANNOTATIONS ARE EVALUATED. ALSO, THE GROUND-TRUTH LABELS ARE FUSED AS AND (t_{US} *and* t_F) AND OR (t_{US} *or* t_F)

input	GT	Acc.	Prec.	Rec.	f1-Score
int	t_{us}	0.722	0.684	0.653	0.660
phase	t_{us}	0.919	0.901	0.929	0.911
comb.	t_{us}	0.923	0.913	0.914	0.914
int	t_f	0.699	0.659	0.651	0.654
phase	t_f	0.923	0.908	0.927	0.916
comb.	t_f	0.926	0.913	0.923	0.918
int	AND	0.748	0.687	0.652	0.662
phase	AND	0.944	0.920	0.953	0.934
comb.	AND	0.937	0.912	0.946	0.926
int	OR	0.683	0.662	0.660	0.661
phase	OR	0.909	0.901	0.912	0.905
comb.	OR	0.893	0.885	0.891	0.888

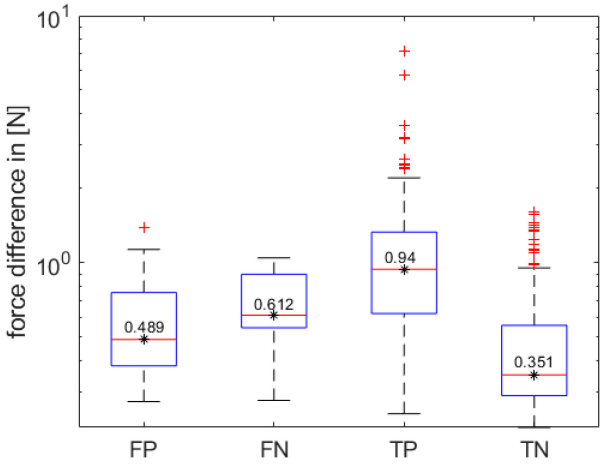

Fig. 6. Boxplot of the prediction capability of our network based on the difference between maximum and minimum force value in the data sample to separate small and large ruptures. The test data was passed to the trained network and the predictions were separated into true positive (TP), true negative (TN), false positive (FP) and false negative (FN) predictions. The TNs median force difference is close to zero and the TPs median is the largest, whilst for the FPs and FNs the median gradients are low. The median values are displayed next to the boxes for all four cases.

best results are achieved with the multi-modal ground-truth label (AND), using phase data as input, with a maximum accuracy of 0.94. We estimated an average time of 4.6ms to process one sample with our CNN approach, using only phase data as input. This enables a real time processing of the OCT data.

We evaluated the cases where our network was not able to predict the occurrence of ruptures or the absence of such. Therefore, we separate the predictions into true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. We rely on our force label as ground-truth modality and phase data as input. Subsequently we determine the difference between the maximum and minimum force of the 224×224 sized data samples to analyze whether the network has difficulties predicting smaller or larger ruptures. Fig. 6 shows the results of our evaluation. Even though the absolute force values are low, the TPs can be distinguished from the FPs, FNs and TNs. The TPs have the highest median concerning the force

TABLE V
VALIDATION OF CNN TRAINED ON THE FOUR PHANTOM TYPES SEPARATELY (SEP.) AND CROSS VALIDATION BETWEEN THE PHANTOM TYPES (CROSS) USING COMBINED OCT INPUT AND LABEL AND

validation	phantom	Acc.	Prec.	Rec.	f1-Score
sep.	A	0.913	0.895	0.935	0.907
	B	0.938	0.868	0.962	0.904
	C	0.949	0.911	0.967	0.934
	D	0.914	0.889	0.910	0.899
cross	A	0.689	0.842	0.526	0.456
	B	0.777	0.594	0.583	0.587
	C	0.621	0.481	0.480	0.480
	D	0.585	0.499	0.499	0.499

TABLE VI
PERFORMANCE OF THE TRAINED CNN ON OCT INPUTS (INTENSITY, PHASE, COMB.) FROM INSERTIONS IN LIVER AND PROSTATE TISSUE WITH t_F GROUND-TRUTH MODALITY. THE MEAN VALUES AND STANDARD DEVIATIONS FOR THE DIFFERENT SAMPLES ARE SHOWN

liver

input	Acc.	Prec.	Rec.	f1-Score
int	0.68 ± 0.13	0.14 ± 0.29	0.44 ± 0.11	0.48 ± 0.11
phase	0.73 ± 0.15	0.41 ± 0.48	0.63 ± 0.17	0.55 ± 0.20
comb.	0.73 ± 0.15	0.63 ± 0.17	0.41 ± 0.48	0.55 ± 0.20

prostate

input	Acc.	Prec.	Rec.	f1-Score
int	0.60 ± 0.17	0.48 ± 0.02	0.19 ± 0.24	0.39 ± 0.06
phase	0.46 ± 0.10	0.51 ± 0.06	0.33 ± 0.26	0.34 ± 0.04
comb.	0.47 ± 0.07	0.52 ± 0.08	0.34 ± 0.26	0.36 ± 0.04

difference and some large differences shown as outliers. The cases predicted as TNs show median force differences close to zero, as there is no rupture and therefore no remarkable change in the force data present in these data samples. The medians for FPs and FNs lie in between these values, representing smaller ruptures in the force data.

The upper part of Table V depicts the prediction results with combined OCT data input for the various types of phantoms (A-D) trained separately (sep.), meaning only phantoms of the same concentrations are used for training and testing. We use the multi-modal label (AND) as it achieved the best scores in the previous evaluation. The obtained scores are comparable to the scores in Table IV. There are only small differences in between the four types of phantoms.

The lower part of Table V depicts the results for a cross validation between the different phantoms (cross). For each of the four types of phantoms, one was left out in the training data and only used for testing. Phantom A for example denotes, that phantoms of type B, C and D were used for training and type A only for testing. While training separately leads to accuracies around 0.9, the cross validation only obtain accuracy scores between 0.65 and 0.77. Furthermore, the precision and recall values are much lower. We performed the same trainings using phase data as input. The results are comparable to the results obtained with combined data as input.

2) Ex-Situ Study: Table VI shows the CNN results for our ex-situ study. As only the force ground-truth is available for this study, we use the network trained with the force label t_f for evaluation. The results are separated into liver and prostate tissue. Both data sets are tested separately, as the tissue structure

varies. The best results are obtained using the OCT phase data of the liver data set as input. Considering the results for the other input data and the prostate tissue, the values are lower and comparable to the phantom cross validation results (Table V). We also evaluate an entire insertion course to predict ruptures with the trained network. The right part of Fig. 5 visualizes the predictions for a prostate sample overlaid on the force course and OCT intensity M-scan and the corresponding histology image. The relevant tissue boundaries marked by a morphologists are highlighted. The network predictions are highlighted with yellow boxes. As we use a sliding window approach to generate the predictions, the boxes do not represent predictions for one input, but rather the predictions of successive samples, whereas the positive predictions are overlapping. Two time periods follow, where ruptures appear. This correlates to the marked tissue boundaries of the hard node in the histology. The left prediction period starts 0.48s prior to the labeled rupture event, the right is delayed by 0.12s.

IV. Discussion

We introduce an experimental setup that is suitable to perform reproducible needle insertions and obtain complex OCT data from the needle tip and multi-modal ground-truth information of rupture events from US and force measurements. Our tissues phantoms cause rupture events of varying magnitudes, whereas a stiff gelatin mixtures lead to larger ruptures (Section III-B). The high standard deviations for rupture events in our phantoms state that besides the large pork-gelatin boundary ruptures also several small rupture events arise during needle insertion.

Calculating the relative tissue motion from OCT phase data and detecting ruptures in the displacement profiles based on optimal scaling factors results in a maximum accuracy of 0.83. Moreover, the precision and recall scores underline the limitation of this threshold based detection method.

Our deep learning methods can be used to predict ruptures in the OCT intensity or phase data. Phase data in all three input channels leads to promising scores and outperforms our results generated with intensity data as input. Considering the complex OCT data, i.e., intensity and phase data (combined), we achieve scores comparably good as the phase data results. Hence, we assume phase data contains more information about rupture events than intensity data. As previous studies underline, the OCT phase data encompasses important information about the tissue mechanics [21], [22] which is valuable for the evaluation of tissue compression and relaxation during rupture events. The differences between our combined approach and using only phase data balance each other over different training runs and evaluations.

Furthermore, our evaluations show that our model can be trained using different ground-truth modalities. Even though we receive a top score for the multi-modal US and force label (AND), t_{us} and t_{f} themselves as labels are equally good. While the AND label contains the necessary information for US and force, the OR label leads to a poorer performance. Due to the fact that some samples only contain the information of a rupture in US but not in force and vice versa, training with the AND label appears to provide the most reliable ground-truth information as the limitations of both labels separately can be overcome in this approach. The investigation on different rupture types, smaller and larger ruptures, shows that the network performs better for larger ruptures. Most false positive and false negative predictions are made for smaller ruptures with a smaller differences between maximum and minimum force value. Training on all phantoms jointly and training on the phantoms separately does not show much difference, whilst excluding a phantom type for a cross validation, it is noticeable that the results do not match the previous ones. Due to the different tissue structures, stiffness (Table III) and varying speckle characteristics in the OCT data, the network seems to be limited to generalize on the unknown phantom types. We expect to obtain better results for our cross validations with a larger data set with more variance.

Summarized, the phantom study underlines that our deep learning methods improve the prediction of rupture events in the complex OCT data of up to 0.12 percentage points compared to the scaling factor based detection method. Our approach to combine two separately trained networks using SVMs and complex OCT data as input (combined) leads to satisfying prediction results with top scores concerning accuracy, precision, recall and f1-score. Average processing times of 4.6ms per sample with an input width of approximately 2.46s, allow for real time processing of the OCT data.

The results of the ex-situ study show that predicting ruptures with our trained CNN is possible but needs improvement. As the outcome of the phantom cross validation shows comparable results, it seems that besides the varying mechanical properties of the ex-situ tissue samples and phantoms the speckle characteristics differ considerably. Furthermore, it could not be observed whether the force ground-truth was reliable in all situations, as no US imaging was possible in the ex-situ study. A further clinical study to record a larger data set of ex-situ tissue samples with additional US guidance could help. The additional data could be used to train the network also on clinical data and thereby improve the adaptation on the unseen tissue data.

We did not evaluate the occurrence of artifacts in the data so far. For example, artifacts induced by tissue motion in the field-of-view of the needle probe could potentially disturb the detection of ruptures. Nevertheless, a network trained on a suitable data set with artifacts represented, should be capable of either compensating or ignoring the occurrence of some artifacts.

Especially for punctures in deeper tissue structures, our deep learning approach using OCT imaging at the needle tip for rupture detection complements external US guidance or force sensing. The complex OCT data offers both morphological and mechanical information of the tissue with a high spatial and temporal resolution. The potential of the latter is not fully exploited in this study as our phantoms and ex-situ tissues contain only a few boundaries with comparatively high distances to each other. In future work, this rupture detection methods will be further investigated to distinguish between different tissue with varying texture and elasticities, e.g. due to a tumor formation, which might be related to lower or higher amount of small or large ruptures. Even small inhomogenities might be detected

employing the so far unused high temporal resolution of the complex OCT data.

V. Conclusion

We present a deep learning approach to detect rupture events in complex OCT data from needle insertions. We use an experimental setup with additional ultrasound imaging and force sensing for ground-truth generation. Our CNN is based on the multi-modal labels and complex OCT data and outperforms conventional approaches based on the estimated relative tissue speed [22]. Applying pretrained CNNs to OCT data recorded at the needle tip, rupture detection is enabled even for deeper tissue structures where US imaging and force measurements are limited. This approach can complement a robust robotic needle placement in the future.

Acknowledgment

The authors acknowledge support for the Open Access fees by Hamburg University of Technology (TUHH) in the funding programme Open Access Publishing. The presented ex-situ study was ethically examined (WF-044/19) by the ethics committee of the medical association Hamburg and stated as uncritical as a backtracking of the analyzed tissue samples to the body donors is not possible.

References

[1] C. A.-J. Webb *et al.*, "Unintentional dural puncture with a tuohy needle increases risk of chronic headache," *Anesth. Analg.*, vol. 115, no. 1, pp. 124–132, 2012.
[2] N. Hollister *et al.*, "Minimising the risk of accidental dural puncture with epidural analgesia for labour: A retrospective review of risk factors," *Int. J. Obstet. Anesth.*, vol. 21, no. 3, pp. 236–241, 2012.
[3] A. Mastmeyer *et al.*, "Evaluation of direct haptic 4 d volume rendering of partially segmented data for liver puncture simulation," *Sci. Rep.*, vol. 7, no. 671, pp. 1–15, 2017.
[4] D. J. van Gerwen *et al.*, "Needle tissue interaction forces: A. survey of experimental data," *Med. Eng. Phys.*, vol. 34, no. 6, pp. 665–680, 2012.
[5] A. M. Okamura *et al.*, "Force modeling for needle insertion into soft tissue," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 10, pp. 1707–1716, Oct. 2004.
[6] S. P. DiMaio and S. E. Salcudean, "Needle insertion modeling and simulation," *IEEE Trans. Robot. Autom: Spec. Issue Med. Robot.*, vol. 19, no. 5, pp. 864–875, Oct. 2003.
[7] I. M. Buzurovic *et al.*, "A novel approach to an automated needle insertion in brachytherapy procedures," *Med. Biol. Eng. Comput.*, vol. 56, no. 2, pp. 273–287, 2018.
[8] Z. Zhou *et al.*, "Optical surgical instrument tracking system based on the principle of stereo vision," *J. Biomed. Opt.*, vol. 22, no. 6, pp. 1–14, 2017.
[9] M. Schlüter *et al.*, "Concept for markerless 6 d tracking employing volumetric optical coherence tomography," *Sensors*, vol. 20, no. 9, p. 2678, 2020.
[10] A. M. Franz *et al.*, "Electromagnetic tracking in medicine-a review of technology, validation, and applications," *IEEE Trans. Med. Imag.*, vol. 33, no. 8, pp. 1702–1725, Aug. 2014.
[11] S. Hebard and G. Hocking, "Echogenic technology can improve needle visibility during ultrasound-guided regional anesthesia," *Regional Anesth. Pain Med.*, vol. 36, no. 2, pp. 185–189, 2011.
[12] W. Xia *et al.*, "Looking beyond the imaging plane: 3 d needle tracking with a linear array ultrasound probe," *Sci. Rep.*, vol. 7, no. 1, p. 3674, 2017.
[13] M. P. Fronheiser *et al.*, "Vibrating interventional device detection using real-time 3-d color doppler," *IEEE Trans. Ultrasonics, Ferroelect. Freq. Control*, vol. 55, no. 6, pp. 1355–1362, Jun. 2008.
[14] S. Elayaperumal *et al.*, "Mr-compatible biopsy needle with enhanced tip force sensing," in *Proc. 2013 World Haptics Conf.*, Daejeon, Korea (South), 2013, pp. 109–114.
[15] Z. Mo *et al.*, "In-vivo tissue identification on mice using a fiber optical tip force sensing needle," *IEEE Sensors J.*, vol. 18, no. 15, pp. 6352–6359, Aug. 2018.
[16] M. Brynolf *et al.*, "Optical detection of the brachial plexus for peripheral nerve blocks: An in vivo swine study," *Regional Anesth. Pain Med.*, vol. 36, no. 4, pp. 350–357, 2011.
[17] J. Desroches *et al.*, "A new method using raman spectroscopy for in vivo targeted brain cancer tissue biopsy," *Sci. Rep.*, vol. 8, no. 1792, pp. 1–10, 2018.
[18] D. Ellebrecht *et al.*, "Towards an optical biopsy during visceral surgical interventions," *Visceral Med.*, vol. 36, no. 2, pp. 70–79, 2020.
[19] B. G. Muller *et al.*, "Needle-based optical coherence tomography for the detection of prostate cancer: A visual and quantitative analysis in 20 patients," *J. Biomed. Opt.*, vol. 23, no. 8, pp. 1–11, 2018.
[20] C. Otte *et al.*, "Investigating recurrent neural networks for oct a-scan based tissue analysis," *Methods Inf. Med.*, vol. 53, no. 4, pp. 245–249, 2014.
[21] S. Latus *et al.*, "An approach for needle based optical coherence elastographymeasurements," *Med. Image Comput. Comput.-Assist. Interv.*, Lecture Notes Comput. Sci., vol. 10434, pp. 655–663, 2017.
[22] K. M. Kennedy *et al.*, "Needle optical coherence elastography for tissue boundary detection," *Opt. Lett.*, vol. 37, no. 12, pp. 2310–2312, 2012.
[23] N. Gessert *et al.*, "Spatio-temporal deep learning models for tip force estimation during needle insertion," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 9, pp. 1485–1493, 2019.
[24] A. G. Roy *et al.*, "Relaynet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Exp.*, vol. 8, no. 8, pp. 3627–3642, 2017.
[25] N. Gessert *et al.*, "Automatic plaque detection in ivoct pullbacks using convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 426–434, Feb. 2019.
[26] R. Rasti *et al.*, "Macular oct classification using a multi-scale convolutional neural network ensemble," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 1024–1034, Apr. 2018.
[27] J. L. Barron *et al.*, "Performance of optical flow techniques," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 43–77, 1994.
[28] K. He *et al.*, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 770–778.
[29] H. Shin *et al.*, "Deep convolutional neural networks for computer aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
[30] J. Deng *et al.*, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, 2009, pp. 248–255.
[31] Y. Ma and G. Guo, *Support Vector Machines Applications*, Berlin, Germany: Springer, Cham, 2014.
[32] A. Krizhevsky *et al.*, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[33] K. Püschel *et al.*, "New developments and possibilities in the field of post-mortem medicine mortui vivos docent," *Rechtsmedizin*, vol. 30, no. 6, pp. 425–429, Aug. 2014.