This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

Development of Convolutional Neural Network to Segment Ultrasound Images of Histotripsy Ablation

Katherine Miao, Katia Flores Basterrechea, Sonia L. Hernandez, Osman S. Ahmed, Mikin V. Patel, Kenneth B. Bader *, *Member, IEEE*

***Abstract*— *Objective:* Histotripsy is a focused ultrasound therapy that ablates tissue via the action of bubble clouds. It is under investigation to treat a number of ailments, including renal tumors. Ultrasound imaging is used to monitor histotripsy, though there remains a lack of definitive imaging metrics to confirm successful treatment outcomes. In this study, a convolutional neural network (CNN) was developed to segment ablation on ultrasound images. *Methods:* A transfer learning approach was used to replace classification layers of the residual network ResNet-18. Inputs to the classification layers were based on imaging assessment of red blood cell phantoms that were ablated by histotripsy ablation, including ultrasound images and digital photographs that served as the ground truth. The efficacy of the CNN was compared to subtraction imaging, and manual segmentation of images by two board-certified radiologists. *Results:* The CNN had a similar performance to manual segmentation, though was improved relative to segmentation with subtraction imaging. Predictions of the network improved over the course of treatment, with the Dice similarity coefficient less than 20% for fewer than 500 applied pulses, but 85% for more than 750 applied pulses. The network was also applied to ultrasound images of *ex vivo* kidney exposed to histotripsy, which indicated a morphological shift in the treatment profile relative to the phantoms. These findings were consistent with histology that confirmed ablation of the targeted tissue. *Conclusion:* Overall, the CNN showed promise as a rapid means to assess outcomes of histotripsy and automate treatment. *Significance:* Data collected in this study indicate integration of CNN image segmentation to gauge outcomes for histotripsy ablation holds promise for automating treatment procedures.**

***Index Terms* — artificial intelligence, convolutional neural network, histotripsy, kidney ablation**

I. Introduction

RENAL cell carcinoma (RCC) is a major burden on the American public, with an estimated incidence of ~ 600k patients in the United States projected for 2023 [1]. Surgery is the gold standard for kidney tumor management in localized cases [2], [3]. The number of patients eligible for surgery has decreased in recent years due to the aging, progressively co-morbid population [4], [5]. Ablation may be an alternative means to address RCC, though existing modalities are invasive, must be applied away from critical renal structures, and have higher recurrence rates compared to surgical management [6]. One potential alternative form of ablation gaining traction is histotripsy, a focused ultrasound therapy that disintegrates tissue via the nucleation and subsequent action of bubble clouds [7], [8]. Ultrasound pulses of more than 25 MPa peak negative pressure and less than 20 μs in duration are applied noninvasively to generate bubble clouds spontaneously within the tissue (i.e., without the need for exogenous agents) [9]. The efficacy of histotripsy for the treatment of renal tumors has been established in pre-clinical studies [10]–[12]. Further, a first-in-human trial is underway to establish safety of the technology for patients with renal tumors (clinicaltrials.gov identifier NCT05820087), and has been cleared for liver lesions.

The degree of histotripsy bubble activity necessary for ablation is not *a priori* known, and will depend on the tissue composition [13]. Histotripsy sources are outfitted with an ultrasound imaging system to confirm accurate targeting and assess treatment outcome [14], [15]. Following treatment, physicians must rely on computed tomography or MRI to confirm outcomes due to the variable appearance of ablation on ultrasound [16]. Additional imaging increases the time, complexity, and cost of the histotripsy procedure [17], making methods to improve the sensitivity and specificity of ultrasound imaging to ablation an active area of research. One potential solution is artificial intelligence, which is rapidly being integrated into medical imaging to aid analysis of pathological features [18]. Residual networks have become state-of-the-art for object recognition tasks [19], improving performance over classical convolutional neural networks [20]. Integration of a neural network into histotripsy image guidance systems will enable automation of the procedure, and allow real-time adjustment of the exposure duration to ensure successful outcomes.

In this study, a residual network was developed to segment ultrasound images of histotripsy ablation. As an initial step to ensure feasibility, the network was trained using *in vitro* data to

This work was supported in part by the American Cancer Society under Grant RSG-21-171-01-ET.

K.M. is with the Department of Biology, University of Chicago, Chicago, IL, USA. K.F.B., O.S.A., and M.V.P. are with the Department of Radiology at the University of Chicago, Chicago, IL, USA. S.L.H. is with the Department of Surgery at the University of Chicago, Chicago, IL, USA. *K.B.B. is with the Department of Radiology, University of Chicago, Chicago, IL, USA (correspondence e-mail: baderk@uchicago.edu).

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

enable accurate registration between ultrasound images and ground truth. The network performance was compared to manual segmentation of images by two interventional radiologists experienced with histotripsy technology. Outcomes for segmentation were also evaluated with subtraction imaging as a rudimentary means to segment images. Further, the network was applied to ultrasound images of *ex vivo* kidney exposed to histotripsy.

II. Methods

A. In Vitro Phantoms and Ex Vivo Kidney

Agarose red blood cell phantoms were produced using an established protocol [21], [22]. Briefly, blood was collected and anticoagulated from female farm raised pigs (40-50 kg). The blood was centrifuged at 3,000 RPM for 10 min to remove the plasma and buffy coat, leaving behind a volume of red blood cells. Low gelling temperature agarose powder (4 g, Sigma-Aldrich, St. Louis, MO) and NaCa (3.6 g, Sigma-Aldrich, St. Louis, MO) were dissolved into 400 mL of 0.2 µm filtered, deionized water by heating in a microwave of 700 W power in 30 second increments until clear. The agarose mixture was transferred to an ultrasonic cleaning bath heated to 55˚C while continuously evacuating (10 kPa) over the course of 30 min. After degassing, ~ 50 mL of the agarose solution was poured into a rectangular acrylic mold and allowed to solidify. A 5% v/v red blood cell/agarose mixture was pipetted onto the solidified agarose slab to form a layer ~ 500 µm thick. After the red blood cell layer solidified, the remainder of the mold was filled with agarose.

Kidney samples were collected from female farm raised pigs (40-50 kg) following euthanasia and used within 72 hrs (N = 8). The samples were sectioned to ~ 2 cm x 2 cm x 2 cm, and embedded in low gelling temperature agarose. All samples (*in vitro* phantoms and *ex vivo* kidney) were submerged in isotonic solution (0.9% NaCl w/v) and degassed (10 kPa) for 2 hours prior to histotripsy exposure.

B. Experimental Set Up

An overview of the experimental set up is shown in **Fig. 1**. An eight element focused source was used to generate bubble clouds. The source had an outer diameter of 10 cm, focal length of 7.5 cm, and fundamental frequency of 1 MHz. Pulses of 20 µs duration (20 acoustic cycles) and 35 MPa peak negative pressure were applied to each target. The source was driven with a custom class-D amplifier and matching network [23]. The focal pressure of the source was measured using a fiber optic hydrophone (HFO-690, Onda Corporation, Sunnyvale CA) up to pressure levels at the threshold for bubble nucleation (~ 25 MPa peak negative pressure [24]). A linear extrapolation between the voltage applied to the amplifier and focal peak negative pressure was used to estimate driving levels beyond the calibration as described previously [25], [26].

The focus of the transducer was aligned to a depth of 2 cm in the sample. Bursts of 20 histotripsy pulses were applied at a rate

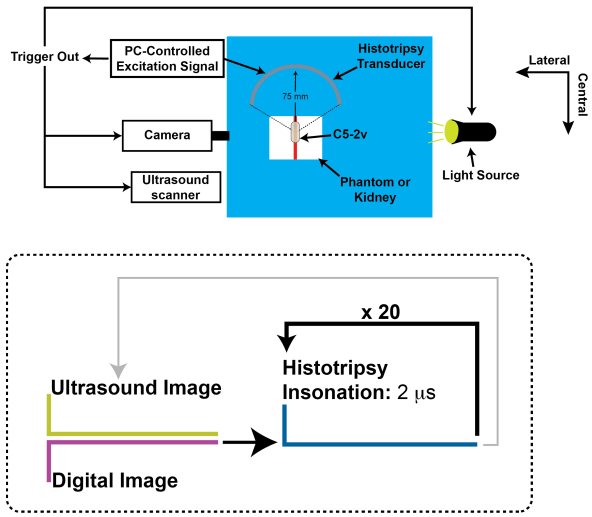

Fig. 1. (Top) Schematic of experimental set up. The coordinate axes refer to the therapy source. The elevational dimension of the therapy source is into the page. (Bottom) Timing diagram for collection of ultrasound imaging data and photographs of phantoms.

of 50 Hz at a fixed location. A total of 2,000 or 4,000 pulses were applied to the *in vitro* phantoms or *ex vivo* tissue, respectively. These treatment durations produced visible ablation on the ultrasound image for each respective target. After each burst of 20 histotripsy pulses was applied, an ultrasound image was acquired with a curvilinear imaging probe (C5-2v, Verasonics, Inc, Kirkland, WA) controlled by a research ultrasound scanner (Vantage 128, Verasonics, Inc, Kirkland, WA). The imaging plane of the probe was along the central/elevational dimension of the therapy source (**Fig. 1**), and the probe to focal zone distance was 60 mm. At this depth, the peak negative pressure of the imaging pulse was 200 kPa (mechanical index of 0.1 [27]).

For the *in vitro* phantoms, a 2.8 MP digital CMOS camera (LUCID Vision Labs, Inc., Elf Place, BC) was also triggered to capture a digital image of the red blood cell phantom after the application of every 20 histotripsy pulses (**Fig. 1**). The camera was fitted with a lens (50 mm focal length, *f* number 2.5 – 16, 1stVision Inc, Andover, MA), resulting in an image resolution of ~ 5 µm per pixel. A pulsed light-emitting diode array (Luma10, HitLights, Chino, CA) was placed opposite the camera for illumination. In total, 100 ultrasound images and digital photographs were acquired for a given *in vitro* dataset. Images were downloaded and processed offline. Digital images of the phantom were segmented using Otsu's method to determine the ablation area (**Fig. 2**). Ultrasound images were upsampled (256 x 1024 pixel dimensions) and co-registered with the digital images using a custom geometric transformation in MATLAB based on hyperechoic nylon filaments embedded within the phantom [28].

C. Training the CNN

Training and evaluation of the convolutional neural network (CNN) was performed on a iMAC (3.2 GHz Quad-Core Intel Core i5, 32 GB memory, OS 12.6.5). An overview of the CNN training is outlined in **Fig. 3**. A total of nine co-registered

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

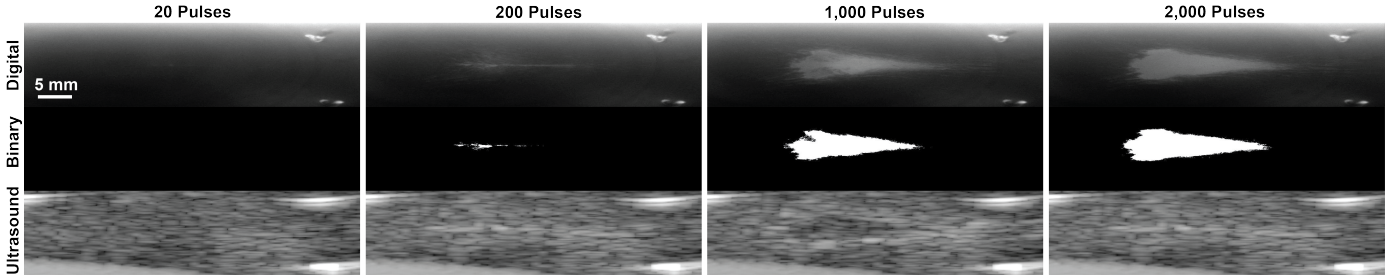

Fig. 2. (Top Row) Photographs of phantom (Middle Row) Segmentation of phantom images. White pixels indicate areas of ablation. (Bottom Row) Ultrasound images of phantom. Hyperintense pixels in the corners of the image correspond to fiducial wires within the phantom used to co-register ultrasound and camera images. The number of applied histotripsy pulses are indicated at the top of each column.

datasets comprised of 900 total images were used to train the CNN to identify ablation in ultrasound images. Training was done in MATLAB (version 2022b, Mathworks, Natick, MA) with the 'trainNetwork' function. The classification layers of the two-dimensional residual network ResNet-18 were replaced with binary classification based on segmentation of the digital images. The data set was split randomly using 60% for training (540 images), 20% for validation (180 images), and 20% for testing (180 images). Images were processed in a randomized order for training, validation, and testing to avoid integrating a bias into the CNN based on stage of treatment. To reduce the negative effect of class imbalance, the Dice loss function was maximized during training. The base learning rate was 0.001, and was decayed using a stochastic gradient descent over 30 epochs. To avoid overfitting, data augmentation was conducted via random reflection and translation of the training data sets. Other hyperparameters were tuned based on a pilot study with a subset of 20 images.

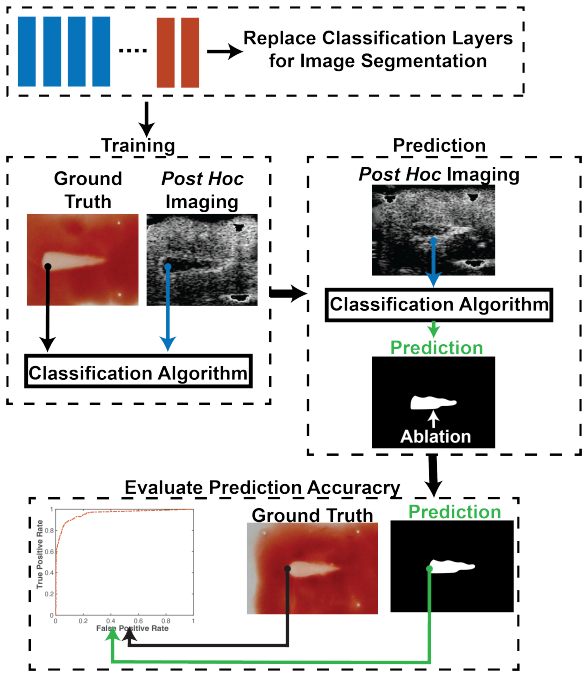

Fig. 3. Overview for training and testing the CNN to segment ablation on ultrasound images.

Table I. Performance metrics of the CNN, subtraction imaging, and manual segmentation to gauge outcomes of histotripsy phantom ablation.

Metric	-----	CNN	Subtraction	Manual
Accuracy	Viable	0.99	0.96	0.99
	Ablated	0.63	0.44	0.64
	Average	0.81	0.70	0.82
IoU Score	Viable	0.95	0.90	0.95
	Ablated	0.58	0.31	0.58
	Average	0.92	0.73	0.92
F1 Score		0.95	0.48	0.73
AUROC		0.92	0.70	0.82
Calculation Time/Image		0.5 ± 0.001 s	0.9 ± 0.8 ms	> 15 s

D. Additional Segmentation Methods

Segmentation was also performed with a subtraction method. Images acquired over the course of histotripsy exposure were subtracted on a pixel-by-pixel basis from a baseline image. Pixels associated with changes in the phantom structure were determined in the subtraction image using Otsu's method (**Fig. 4**). A subset of data were manually segmented by two board-certified interventional radiologists, each with over ten years of experience interpreting clinical images. The interventional radiologists were informed that images with and without ablation were included. The images were analyzed in a randomized order to limit potential biases in segmentation. A total of 132 images (44 *in vitro* and 88 *ex vivo*) were manually segmented with a custom script in MATLAB.

E. Quantification of Segmentation

Each segmentation method (CNN, subtraction imaging, and manual segmentation) was evaluated on four *in vitro* datasets comprised of 400 images that were not included in the training datasets. For each segmentation type, the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) pixels were tabulated to report the accuracy:

$$\text{Accuracy} = \frac{\text{TP}}{\text{TP+FN}} \quad (1)$$

and Dice Similarity Coefficient (DSC):

$$\text{DSC} = \frac{2\text{TP}}{\text{A}_{\text{SEG}} + \text{A}_{\text{Truth}}} \quad (2)$$

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

where A is area, and the subscripts indicate either segmentation (*SEG*) or truth (*Truth*). Over or under segmentation of the predictions were assessed with the IoU Score:

$$\text{IoU} = \frac{\text{TP}}{\text{TP+FP+FN}} \quad (3)$$

The ability of each segmentation method to identify the ablation zone contours was evaluated with the boundary F1 (BF) matching score, based on the precision (TP/ TP + FP) and recall (TP/TP + FN) values as [29]:

$$\text{BF Score} = \frac{2 \text{ x precision x recall}}{\text{precision + recall}} \quad (4)$$

The area under the receiver operator characteristics curve (AUC) was calculated to evaluated segmentation performance from contingency tables as [30]:

$$\text{AUC} = \frac{\text{TP}}{2(\text{TP+FN})} + \frac{\text{TN}}{2(\text{FP+TN})} \quad (5)$$

The maximum and mean Hausdorff distance h were also computed to gauge the distance between the predicted and actual ablation areas [31]:

$$h\,(A_{CNN}, A_{Truth}) = \max_{a \in A_{SEG}} \text{or mean} \{ \min_{b \in A_{Truth}} \{d(a, b)\}\} \quad (6)$$

where a/b are spatial locations of segmented/truth, and $d(a,b)$ is the distance between these points.

F. Evaluation of Ex Vivo Kidney

Imaging data collected during histotripy insonation of the *ex vivo* samples were processed with the CNN, subtraction imaging, and manual segmentation. To confirm ablation, kidney samples exposed to histotripsy were fixed for 24 – 36 hrs in 10% buffered formalin (Sigma-Aldrich, St. Louis, MO), transferred to 70% reagent alcohol (Thermo Scientific, Waltham, MA) for at least 72 hrs, and embedded in paraffin. Embedded tissues were sectioned to 5 µm thickness, and stained with hematoxylin and eosin (H&E). Stained tissues were then scanned at 20x magnification (Vs200, Olympus Life Sciences, Waltham, MA), and analyzed in QuPath [32].

III. Results

A. Overall Performance of CNN

A total of 1350 iterations were used to train the network, which had a final validation loss of 3.4% and accuracy of 98.7%. The CNN required 0.5 ± 0.001 seconds per image for segmentation on an iMAC computer (3.2 GHz Quad-Core Intel Core i5, 32 GB memory, OS 12.6.5). Outcomes for accuracy, IoU Score, Boundary F1 Score, and AUROC over all 400 testing images are indicated in **Table I**. The increase in AUROC relative to 0.5 indicated that the network did better than guessing in terms of predicting the ablation area ($p < 0.05$) [30]. Outcomes for the CNN were similar to those based on manual segmentation, but improved relative to subtraction imaging (**Table I**).

Representative outcomes for the CNN evaluation of *in vitro* phantom datasets are shown in **Fig. 5**. There was good qualitative agreement between the CNN and truth during the later stages of treatment. The CNN tended to underpredict the ablation extent during the early stages of treatment. Interestingly, the appearance of the ablation zone changed over the course of histotripsy exposure from hyperintense to hypointense (**Fig. 6**). This change may have been a contributing factor to the observed behavior of the CNN.

B. CNN Performance Over the Course of Histotripsy Exposure

The ablation area relative to the number of applied histotripsy pulses, and predictions for each segmentation method are shown in **Fig. 7**. There was a rapid increase in the ablation area over the course of ~ 600 applied pulses, after which the rate of growth declined. A similar behavior was observed for the CNN, though the onset of rapid grow was delayed until the application of ~ 500 pulses. Predictions of the CNN coincided with the actual ablation area after ~ 750 pulses were applied ($p = 0.21$). Little to no agreement was observed between areas identified in subtraction imaging and the truth (**Fig. 7**).

Similar to the automated segmentation algorithms, manual segmentation generally underestimated the ablation area. Between two and four total false positives (i.e., no ablation identified when ablation was present) were noted at each treatment duration evaluated by the observers. The presence of total false positives increased the standard deviation for manual segmentation relative to the automated methods.

The histotripsy pulse dependence for accuracy, Dice Similarity Coefficient, and Hausdorff distance were also computed over the course of histotripsy exposure (**Fig. 8**). These parameters indicated the CNN had maximum efficacy when images were acquired after more than ~ 750 histotripsy pulses were applied. For subtraction imaging, the accuracy, DSC, and mean Hausdorff distance also had positive trends with the number of applied pulses, though were reduced relative to the CNN. The Hausdorff distances (both mean and

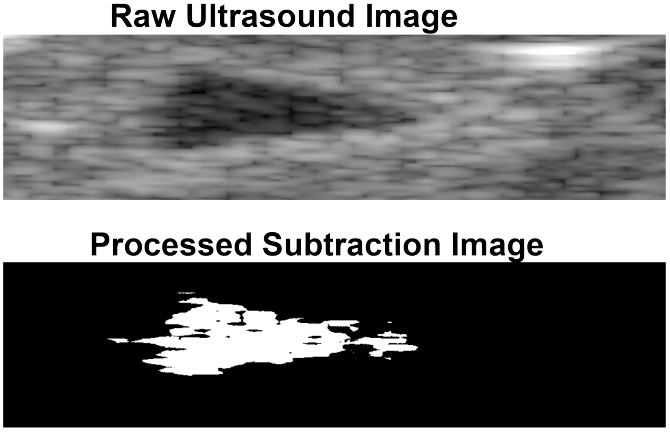

Fig. 4. (Top) Ultrasound image of phantom ablation after 2,000 applied pulses. (Bottom) Corresponding subtraction image. White pixels are considered changes in echogenicity associated with ablation.

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

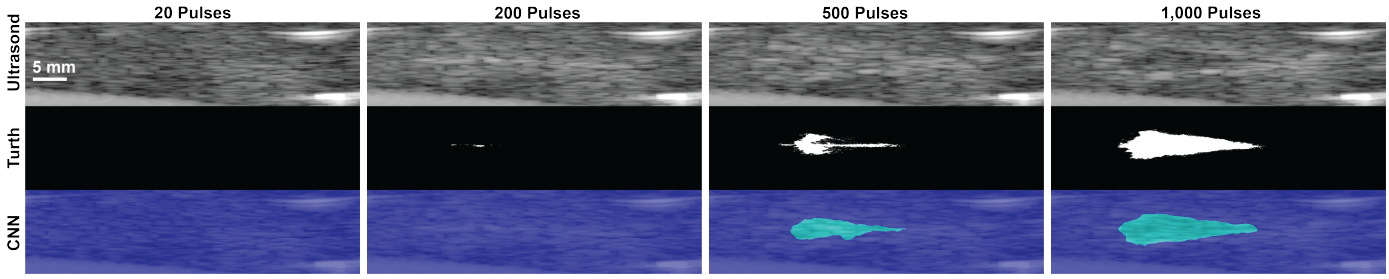

Fig. 5. (Top Row) Ultrasound images of phantom ablation. (Middle Row) Truth images indicating locations of phantom ablation (white). (Bottom Row) Overlay of ultrasound images with predictions of ablation with the CNN (turquoise). The number of applied histotripsy pulses are indicated at the top of each column.

maximum) were improved for manual segmentation relative to the automated algorithms for ~ 250 to 550 applied pulses. Note that total false negatives were not be included in the analysis for Hausdorff distance (i.e., there is no '*a*' term as input to **(6)**). These data indicate observers were able identify the ablation zone effectively when apparent on the image, though were less able to delineate subtle indications of ablation.

C. *Outcomes for Ex vivo Kidney*

Representative images of *ex vivo* kidney tissue over the course of histotripsy exposure are shown in **Fig. 9**. In contrast to the *in vitro* phantoms, targeted regions exhibited a general reduction in pixel intensity over the course of treatment for the *ex vivo* tissue. Regions of hypoechogenicity were generally well identified by the CNN and subtraction imaging, as noted in **Fig. 9**. The regions noted in subtraction imaging were generally larger than those outline by the CNN.

Outcomes for all the segmentation methods over the course of histotripsy exposure are indicated in **Fig. 10**. Areas identified on subtraction imaging increased rapidly up to the application of 630 ± 395 pulses, after which there was little change. In contrast, areas identified by the CNN increased continuously over the course of histotripsy exposure. Outcomes for the CNN and subtraction imaging coincided after ~ 1280 pulses were applied ($p = 0.09$).

Results for manual segmentation by two interventional radiologists were mixed. One interventional radiologist generated false positives for all negative control images (i.e., identified ablation in control images). Therefore, all *ex vivo* data generated by this interventional radiologist was excluded from analysis. Outcomes from the remaining interventional radiologist were consistent with outcomes for subtraction imaging and the CNN. All three significantly correlated with the treatment duration ($p < 0.05$). The Pearson correlation coefficient for the observer and CNN were nearly identical (~ 0.97). In contrast, the correlation coefficient for subtraction imaging was ~0.56.

Histotripsy ablation was apparent on gross analysis of kidney samples (**Fig. 11**, Left Panel), with minimal structure remaining intact. Non-targeted kidney tissue had cortex structure consisting of intact glomeruli and tubule structures (**Fig. 11**, Middle Panel). Over a 100 µm distance into targeted portions of the sample, there was a loss of organization or defining features of glomeruli and tubules. Further, the treatment zone was devoid of nuclei, and contained interspersed regions without cellular structures (i.e., loss of hematoxylin and eosin). Blood vessels larger than ~ 100 µm in diameter were largely unaffected (**Fig. 11**, Right Panel).

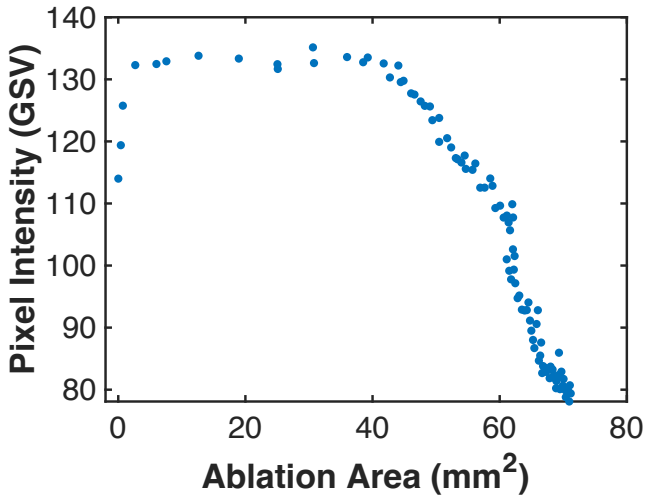

Fig. 6. Representative change in the ultrasound pixel intensity for regions associated with ablation relative to the size of the ablation zone. During the early stage of insonation, ablated pixels appeared hyperintense on ultrasound images. After the ablation area has grown sufficiently large, the ablation area appeared hypointense.

IV. DISCUSSION AND CONCLUSIONS

A. *Outcomes of the CNN*

Histotripsy is a promising approach to address limitations with current treatments for renal tumors in the adult and pediatric patient populations [12], [33]. The development of quantitative imaging markers to assess treatment outcomes will aid in the widespread clinical adoption of histotripsy. Multiple markers are under development for histotripsy, including passive imaging based on acoustic emissions generated by bubble activity [28], [34] and contrast-enhanced ultrasound

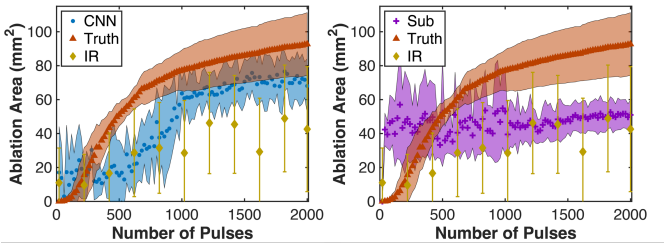

Fig. 7. (Left) The pulse dependence for the ablation area (Truth) and prediction of the CNN relative to the number of applied histotripsy pulses. (Right) The pulse dependence for the ablation area (Truth) and prediction of subtraction imaging to the number of applied histotripsy pulses. Diamonds represent the mean and standard deviation for manual segmentation of images by interventional radiologists. Solid markers indicate the mean (N = 4) and shaded areas (bars for manual segmentation) indicate the standard deviation.

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

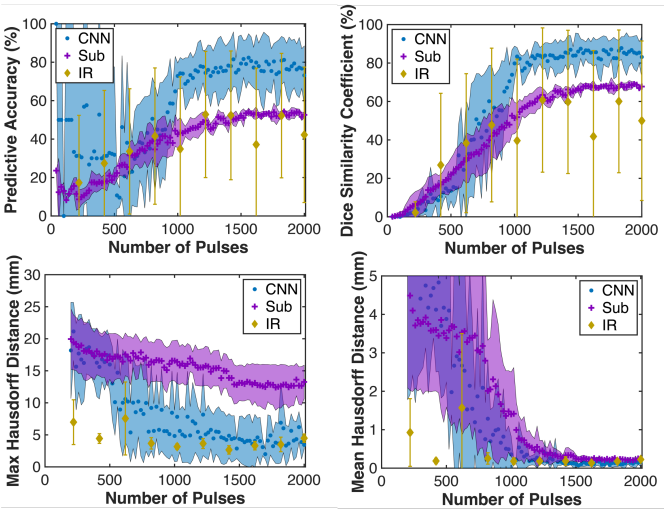

Fig. 8. Assessment of the histotripsy pulse-dependence for the accuracy, Dice Similarity Coefficient, and Hausdorff distance (maximum and mean). Analysis was limited for datasets collected when less than ~ 250 applied pulses due to little to no actual ablation. solid markers indicate the mean (N = 4) and shaded areas (bars for manual segmentation) indicate the standard deviation.

imaging [35]. Acoustic emission correlate with treatment outcomes, though are indirectly related to changes in tissue structure. Contrast imaging delineates non-perfusing regions that correlate with the ablated area. Histotripsy pulses cannot be applied concurrently with contrast imaging due to the administration of microbubbles, limiting its utility to provide real-time feedback during treatment.

Based on the success of artificial intelligence to segment medical images [36], this study developed a convolutional neural network (CNN) to rapidly segment non-contrast ultrasound images of histotripsy ablation. Overall, the CNN provided reasonable assessment of outcomes, with improved performance compared to an image subtraction method, and similar outcomes to manual segmentation by trained interventional radiologists (**Table I**). Interestingly, the AUROC and F1 scores were improved for the CNN relative to manual segmentation, which indicates the algorithm is adept at identifying the ablation zone boundary [29], and therefore therapy margins.

Analysis based on the number of applied pulses provided additional insights for each segmentation method (**Fig. 7** and **8**). The CNN was accurate to within 20% of the ablation area during the latter stages of treatment (> 750 applied pulses), consistent with "successful" imaging feedback outcomes for other ablative therapies [37]. The CNN consistently underestimated the ablation zone. Interestingly, this suggests there are some features of ablation that are not fully captured by the ultrasound image. There may be means to extract this information through multimodal ultrasound imaging, though it may be more difficult to execute such a sequence in real time. Nevertheless, the underestimation of the ablation area should be a consideration when overtreatment margins are included in the therapy protocol [38].

In contrast, the accuracy of the CNN was significantly reduced over the first 750 applied pulses, which may impact what aspects of treatment can be monitored. In addition to ablation, histotripsy has been shown to promote an antitumor immune response in pre-clinical studies [39] and in a case report for in-human liver tumor ablation [40]. Exposures that do not fully ablate the target appear to be critical for inducing an immune response [41], and may not be accurately capture by the network developed here. The poor performance of the CNN early in the treatment may be due to the shifts in the appearance of ablation on ultrasound imaging (**Fig. 6**). Future work may aggregate multiple networks to address the change in ablation zone appearance over the course of treatment, or integrate time-dependent information for training.

B. Utility of the CNN

A particular advantage of the CNN is its speed relative to manual segmentation (**Table I**). Future work will integrate GPU computing to execute the CNN, which is estimated to increase the processing speed to ~ 20 Hz (~ 50 ms interframe time). Histotripsy pulses can be applied at rates up to 1,000 Hz [42], though are generally applied at ~ 100 Hz [43]. Evaluation of the CNN at 20 Hz would therefore assess information over the application of every ~ 4 applied pulses. In this study, the ablation area increased at most ~ 5% over the course of four applied pulses.

One confounding factor for real-time implementation is the influence of hyperechoic bubble clouds within the focal zone. Histotripsy bubble clouds can persist several seconds following insonation [44], whereas images analyzed here were collected > 10 s after histotripsy exposure to assess changes to the target. Contrast-specific imaging may provide a means to remove contribution of bubble activity prior to analysis with the CNN [45]. Further, bubble-deleting pulses could be applied to remove hyperechoic bubbles during image acquisition [46].

C. Analysis of Ex Vivo Kidney

The CNN was also applied to data collected from *ex vivo* porcine kidney. While ablation was confirmed following histotripsy exposure (**Fig. 11**), there was no ground truth for segmentation of the *ex vivo* samples due distortion of the tissue in preparation for staining and sectioning. The areas identified by the CNN increased slowly over the course of treatment, and were overall smaller than the areas identified *in vitro*. Further, areas identified by the CNN were 60% ± 28% smaller than those identified via manual segmented. Tissue effects may have affected outcomes of the CNN. The tissue sample has an increased attenuation and aberration relative to the agarose sample [47], which may affect visualization of the ablation zone and therefore the CNN. The phantom was a uniform media, whereas the tissue is heterogenous. Future studies will determine the influence of tissue structure on predictions of the CNN. It should be noted that outcomes *in vivo* will be affected by a reduction in image quality due to the combined effects of attenuation and aberration. Deep seated targets may be particularly challenging due to loss of signal and elevational scatter. For these cases, alternative imaging sequences

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

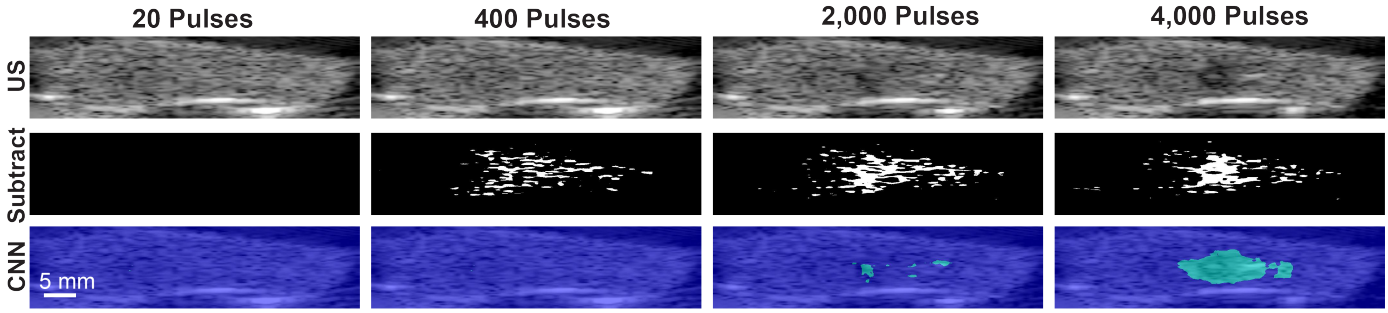

Fig. 9. (Top Row) Ultrasound images of kidney being exposed to histotripsy. Outcomes for subtraction imaging (Middle Row), and the CNN (Bottom Row) are also shown. The number of applied pulses are identified for each column at the top of the image.

developed for deep targets may be of interest, such as chirp-coded excitation [45].

Outcomes for manual segmentation were mixed and observer dependent. The interventional radiologists that performed manual segmentation were informed control images with and without ablation were included in the datasets, and that the image order was randomized to prevent potential biases in segmentation. One radiologist identified false positive ablation in all control images, and their analysis was excluded at all time points (i.e., control and ablation images) for the *ex vivo* datasets. First-in-human trials have highlighted the difficulty of identifying histotripsy ablation with standard B-mode ultrasound for highly trained radiologists [35]. For the remaining observer, no total false negatives were observed in the final stages of treatment (i.e., 4,000 applied pulses). In contrast, observers assigned total false positives at all stages of treatment for the *in vitro* datasets (**Fig. 7**). While the *in vitro* phantoms mimic the appearance of tissue [21], it likely differs from common image features of real tissue evaluated by the observers.

D. Limitations

There are several limitations to this study that prohibit the generalizability of these findings. Here, the network was trained on *in vitro* phantoms, and tested on phantoms and *ex vivo* kidney. Fiducial markers enable accurate registration of the ultrasound images and phantom images [28]. These datasets represent a best case scenario, though may not reflect the CNN performance *in vivo* where bleeding and local inflammation due to ablation may confound strict delineation of outcomes [48]. Translation of this CNN *in vivo* will be a primary focus for future studies based on the information gathered in these *in vitro* and *ex vivo* datasets. To enable collection of both ultrasound and digital images, the imaging probe was placed perpendicular to the therapy source (**Fig. 1**). In practice, imaging probes used for treatment guidance are positioned in a coaxial opening within the therapy source [14], and visualize the target in a different plane. The ablation zone dimensions are nominally symmetric along the lateral and elevational dimensions of the focused transducer [49], and the network could be retrained by adjusting the orientation of the images. A transfer learning approach was used with a ResNet-18 architecture [36]. Residual networks are state-of-the-art for object recognition tasks [19], with an improved performance over plain neural networks and traditional segmentation algorithms [20]. Future studies will evaluate the utility of training the network from scratch rather than a transfer learning approach. Here, the network was trained on images of ablation at a single location, whereas in practice a volume of tissue will be treated consisting of multiple overlapping treatment locations. The CNN is anticipated to have a similar performance volumetric lesions based on their similar hypoechoic appearance to the single focal lesions used for training [7], [50]. Future studies will investigate the CNN accuracy based on the lesion extent. The accuracy of the CNN was not investigated for ablation zones generated in different locations in the imaging plane. The distance between the imaging probe and ablation zone were fixed. Motion under respiration will cause variation in this distance *in vivo*, thereby altering the ablation location within the imaging plane. There are multiple methods that can be used to compensation for respiratory motion, including respiratory gating to trigger

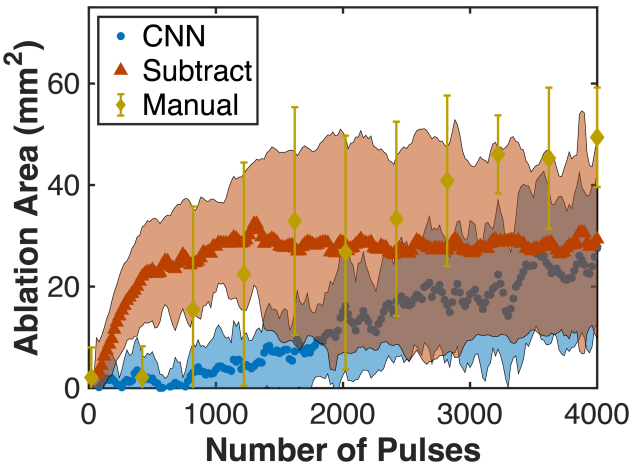

Fig. 10. The histotripsy pulse dependence for areas identified by the CNN, subtraction imaging, and manual segmentation (IR1) during kidney sample exposure. Solid markers indicate the mean (N = 8) and shaded areas (bars for manual segmentation) indicate the standard deviation.

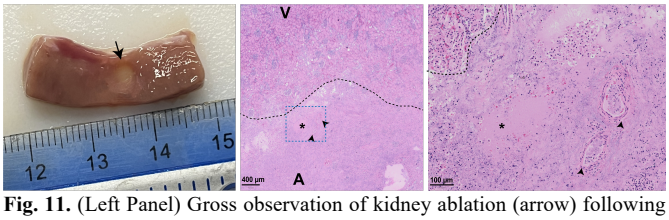

Fig. 11. (Left Panel) Gross observation of kidney ablation (arrow) following histotripsy exposure. (Middle Panel) Hematoxylin and eosin stains indicated viable ('V' above dashed line) kidney cortex comprised of tubules and glomeruli. In contrast, analysis of the ablation zone ('A' below dashed line) revealed absence of organized kidney cortex structures (e.g., removal of tubules and glomeruli), and few disorganized nuclei (*). (Right Panel) Higher magnification insert of blue box in middle image, highlighting viable vessels (arrowheads) and loss of nuclei (*) within the treatment zone.

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

image acquisition [51], active correction techniques under development for histotripsy [52], or high frequency jet ventilation [53]. Nevertheless, these data indicate the use of a CNN for segmenting images to gauge outcomes for histotripsy ablation holds promise for automating treatment procedures.

DATASHARING

Data in this project can be accessed via: https://figshare.com/projects/Machine_Learning_Code/174252

REFERENCES

[1] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023.," *CA. Cancer J. Clin.*, vol. 73, no. 1, pp. 17–48, 2023, doi: 10.3322/caac.21763.

[2] D. C. Miller, J. M. Hollingsworth, K. S. Hafez, S. Daignault, and B. K. Hollenbeck, "Partial nephrectomy for small renal masses: An emerging quality of care concern?," *J. Urol.*, vol. 175, no. 3, pp. 853–858, 2006, doi: 10.1016/S0022-5347(05)00422-2.

[3] B. M. Shinder *et al.*, "Surgical management of advanced and metastatic renal cell carcinoma: A multidisciplinary approach," *Front. Oncol.*, vol. 7, no. MAY, pp. 1–14, 2017, doi: 10.3389/fonc.2017.00107.

[4] D. A. Berger *et al.*, "Impact of Comorbidity on Overall Survival in Patients Surgically Treated for Renal Cell Carcinoma," *Urology*, vol. 72, no. 2, pp. 359–363, 2008, doi: 10.1016/j.urology.2008.02.061.

[5] B. Gershman *et al.*, "Comprehensive Characterization of the Perioperative Morbidity of Cytoreductive Nephrectomy," *Eur. Urol.*, vol. 69, no. 1, pp. 84–91, 2016, doi: 10.1016/j.eururo.2015.05.022.

[6] S. P. Psutka, A. S. Feldman, W. S. McDougal, F. J. McGovern, P. Mueller, and D. A. Gervais, "Long-term oncologic outcomes after radiofrequency ablation for T1 renal cell carcinoma," *Eur. Urol.*, vol. 63, no. 3, pp. 486–492, 2013, doi: 10.1016/j.eururo.2012.08.062.

[7] Z. Xu, T. L. Hall, E. Vlaisavljevich, and F. T. Lee, "Histotripsy: the first noninvasive, non-ionizing, non-thermal ablation technique based on ultrasound," *Int. J. Hyperth.*, vol. 38, no. 1, pp. 561–575, 2021, doi: 10.1080/02656736.2021.1905189.

[8] V. A. Khokhlova *et al.*, "Histotripsy methods in mechanical disintegration of tissue: Towards clinical applications," *Int. J. Hyperth.*, vol. 31, no. 2, pp. 145–162, Mar. 2015.

[9] K. B. Bader, E. Vlaisavljevich, and A. D. Maxwell, "For Whom the Bubble Grows: Physical Principles of Bubble Nucleation and Dynamics in Histotripsy Ultrasound Therapy," *Ultrasound Med. Biol.*, vol. 45, no. 5, 2019, doi: 10.1016/j.ultrasmedbio.2018.10.035.

[10] T. L. Hall, K. Kieran, K. Ives, J. B. Fowlkes, C. A. Cain, and W. W. Roberts, "Histotripsy of Rabbit Renal Tissue in Vivo: Temporal Histologic Trends," *J. Endourol.*, vol. 21, no. 10, pp. 1159–1166, Oct. 2007.

[11] S. C. Mauch *et al.*, "Hepatic and Renal Histotripsy in an Anticoagulated Porcine Model," *J. Vasc. Interv. Radiol.*, vol. 34, no. 3, pp. 386-394.e2, 2023, doi: 10.1016/j.jvir.2022.11.034.

[12] I. Iwanicki *et al.*, "Histotripsy induces apoptosis and reduces hypoxia in a neuroblastoma xenograft model," *Int. J. Hyperth.*, vol. 40, no. 1, p. , 2023, doi: 10.1080/02656736.2023.2222941.

[13] E. Vlaisavljevich, Y. Kim, G. Owens, W. Roberts, C. Cain, and Z. Xu, "Effects of tissue mechanical properties on susceptibility to histotripsy-induced tissue damage," *Phys. Med. Biol.*, vol. 59, no. 2, pp. 253–270, Dec. 2013.

[14] A. D. Maxwell, K. J. Haworth, C. K. Holland, S. A. Hendley, W. Kreider, and K. B. Bader, "Design and Characterization of an Ultrasound Transducer for Combined Histotripsy-Thrombolytic Therapy," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 69, no. 1, pp. 156–165, Jan. 2022, doi: 10.1109/TUFFC.2021.3113635.

[15] T. Hall, J. Fowlkes, and C. Cain, "A real-time measure of cavitation induced tissue disruption by ultrasound imaging backscatter reduction," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 54, no. 3, pp. 569–575, Mar. 2007.

[16] O. Rouvière, R. Souchon, R. Salomir, A. Gelet, J. Y. Chapelon, and D. Lyonnet, "Transrectal high-intensity focused ultrasound ablation of prostate cancer: Effective treatment requiring accurate imaging," *Eur. J. Radiol.*, vol. 63, no. 3, pp. 317–327, Sep. 2007.

[17] S. Gyftopoulos, K. E. Guja, N. Subhas, M. S. Virk, and H. T. Gold, "Cost-effectiveness of magnetic resonance imaging versus ultrasound for the detection of symptomatic full-thickness supraspinatus tendon tears.," *J. shoulder Elb. Surg.*, vol. 26, no. 12, pp. 2067–2077, Dec. 2017, doi: 10.1016/j.jse.2017.07.012.

[18] M. L. Giger, "Machine Learning in Medical Imaging," *J. Am. Coll. Radiol.*, vol. 15, no. Part B, pp. 512–520, Mar. 2018.

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

[19] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," in *Computer Vision -- ECCV 2018*, 2018, pp. 833–851.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[21] A. D. Maxwell, T. Y. Wang, L. Yuan, A. P. Duryea, and Z. Xu, "A tissue phantom for visualization and measurement of ultrasound-induced cavitation damage," *Ultrasound Med. Biol.*, vol. 36, no. 12, pp. 2132–2143, 2010.

[22] S. A. Hendley, V. Bollen, G. J. Anthony, J. D. Paul, and K. B. Bader, "In vitro assessment of stiffness-dependent histotripsy bubble cloud activity in gel phantoms and blood clots," *Phys. Med. Biol.*, vol. 64, no. 14, p. 145019, Jul. 2019, doi: 10.1088/1361-6560/ab25a6.

[23] A. D. Maxwell *et al.*, "A Prototype Therapy System for Transcutaneous Application of Boiling Histotripsy," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 64, no. 10, pp. 1542–1557, Oct. 2017.

[24] A. D. Maxwell, C. A. Cain, T. L. Hall, J. B. Fowlkes, and Z. Xu, "Probability of Cavitation for Single Ultrasound Pulses Applied to Tissues and Tissue-Mimicking Materials," *Ultrasound Med. Biol.*, vol. 39, no. 3, pp. 449–465, Mar. 2013.

[25] Y. Kim, A. D. Maxwell, T. L. Hall, Z. Xu, K.-W. Lin, and C. A. Cain, "Rapid prototyping fabrication of focused ultrasound transducers," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 61, no. 9, pp. 1559–1574, Jan. 2014.

[26] K.-W. Lin, T. L. Hall, R. J. McGough, Z. Xu, and C. A. Cain, "Synthesis of monopolar ultrasound pulses for therapy: The frequency-compounding transducer," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 61, no. 7, pp. 1123–1136, Jul. 2014.

[27] R. E. Apfel and C. K. Holland, "Gauging the likelihood of cavitation from short-pulse, low-duty cycle diagnostic ultrasound," *Ultrasound Med. Biol.*, vol. 17, no. 2, pp. 179–185, 1991.

[28] K. B. Bader, K. J. Haworth, A. D. Maxwell, and C. K. Holland, "Post Hoc Analysis of Passive Cavitation Imaging for Classification of Histotripsy-Induced Liquefaction in Vitro," *IEEE Trans. Med. Imaging*, vol. 37, no. 1, pp. 106–115, 2018, doi: 10.1109/TMI.2017.2735238.

[29] G. Csurka, D. Larlus, and F. Perronnin, "What is a good evaluation measure for semantic segmentation?," in *Procedings of the British Machine Vision Conference 2013*, 2013, pp. 32.1-32.11, doi: 10.5244/C.27.32.

[30] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.," *Biometrics*, vol. 44, no. 3, pp. 837–845, Sep. 1988.

[31] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, 1993, doi: 10.1109/CVPR.1992.223209.

[32] P. Bankhead *et al.*, "QuPath: Open source software for digital pathology image analysis," *Sci. Rep.*, vol. 7, no. 1, pp. 1–7, 2017, doi: 10.1038/s41598-017-17204-5.

[33] N. R. Styn, J. C. Wheat, T. L. Hall, and W. W. Roberts, "Histotripsy of VX-2 Tumor Implanted in a Renal Rabbit Model," *J. Endourol.*, vol. 24, no. 7, pp. 1145–1150, 2010.

[34] J. J. Macoskey, J. R. Sukovich, T. L. Hall, C. A. Cain, and Z. Xu, "Real-time acoustic-based feedback for histotripsy therapy," *J. Acoust. Soc. Am.*, vol. 141, no. 5, p. 3551, 2017.

[35] X. Serres-Creixams *et al.*, "Contrast-Enhanced Ultrasound: A Useful Tool to Study and Monitor Hepatic Tumors Treated With Histotripsy," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 68, no. 9, pp. 2853–2860, Sep. 2021, doi: 10.1109/TUFFC.2021.3073540.

[36] Y. Chandola, J. Virmani, H. S. Bhadauria, and P. Kumar, "End-to-end pre-trained CNN-based computer-aided classification system design for chest radiographs," *Deep Learn. Chest Radiogr.*, pp. 117–140, 2021, doi: 10.1016/b978-0-323-90184-0.00011-4.

[37] M. Anttinen *et al.*, "Feasibility of MRI-guided transurethral ultrasound for lesion-targeted ablation of prostate cancer," *Scand. J. Urol.*, vol. 53, no. 5, pp. 295–302, 2019, doi: 10.1080/21681805.2019.1660707.

[38] M. Morrow *et al.*, "Society of Surgical Oncology–American Society for Radiation Oncology–American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Ductal Carcinoma In Situ," *J. Clin. Oncol.*, vol. 34, no. 33, pp. 4040–4046, Nov. 2016, doi: 10.1200/JCO.2016.68.3573.

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

This article has been accepted for publication in IEEE Transactions on Biomedical Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBME.2024.3352538

[39] S. Qu *et al.*, "Non-thermal histotripsy tumor ablation promotes abscopal immune responses that enhance cancer immunotherapy," *J. Immunother. Cancer*, vol. 8, no. 1, pp. 1–12, 2020, doi: 10.1136/jitc-2019-000200.

[40] J. Vidal-Jove, X. Serres-Creixams, T. Ziemlewicz, and J. M. Cannata, "Liver Histotripsy Mediated Abscopal Effect. Case Report.," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 68, no. 9, pp. 3001–3005, 2021, doi: 10.1109/TUFFC.2021.3100267.

[41] A. L. Pepple *et al.*, "Spatiotemporal local and abscopal cell death and immune responses to histotripsy focused ultrasound tumor ablation," *Front. Immunol.*, vol. 14, no. January, pp. 1–16, 2023, doi: 10.3389/fimmu.2023.1012799.

[42] A. D. Maxwell, G. Owens, H. S. Gurm, K. Ives, D. D. Myers Jr, and Z. Xu, "Noninvasive Treatment of Deep Venous Thrombosis Using Pulsed Ultrasound Cavitation Therapy (Histotripsy) in a Porcine Model," *J. Vasc. Interv. Radiol.*, vol. 22, no. 3, pp. 369–377, Mar. 2011.

[43] A. Shi *et al.*, "Integrated Histotripsy and Bubble Coalescence Transducer for Thrombolysis," *Ultrasound Med. Biol.*, vol. 44, no. 12, pp. 2697–2709, 2018, doi: 10.1016/j.ultrasmedbio.2018.08.013.

[44] E. L. Wallach, H. Shekhar, F. Flores-Guzman, S. L. Hernandez, and K. B. Bader, "Histotripsy Bubble Cloud Contrast With Chirp-Coded Excitation in Preclinical Models," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 69, no. 2, pp. 787–794, 2022, doi: 10.1109/TUFFC.2021.3125922.

[45] V. V. Trivedi, E. L. Wallach, K. B. Bader, and H. Shekhar, "Contrast-Enhanced Imaging of Histotripsy Bubble Clouds Using Chirp-Coded Excitation and Volterra Filtering," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 70, no. 9, pp. 989–998, Sep. 2023, doi: 10.1109/TUFFC.2023.3289918.

[46] A. Shi, Z. Xu, J. Lundt, H. A. Tamaddoni, T. Worlikar, and T. L. Hall, "Integrated Histotripsy and Bubble Coalescence Transducer for Rapid Tissue Ablation," *Ultrason. Ferroelectr. Freq. Control. IEEE Trans.*, vol. 65, no. 10, pp. 1822–1831, Jan. 2018.

[47] K. B. Bader, M. J. Crowe, J. L. Raymond, and C. K. Holland, "Effect of frequency-dependent attenuation on predicted histotripsy waveforms in tissue-mimicking phantoms," *Ultrasound Med. Biol.*, vol. 42, no. 7, pp. 1701–1705, 2016, doi: 10.1016/j.ultrasmedbio.2016.02.010.

[48] M. Burtnyk, T. Hill, H. Cadieux-Pitre, and I. Welch, "Magnetic Resonance Image Guided Transurethral Ultrasound Prostate Ablation: A Preclinical Safety and Feasibility Study with 28-Day Followup," *J. Urol.*, vol. 193, no. 5, pp. 1669–1675, May 2015.

[49] K. B. Bader, K. Flores Basterrechea, and S. A. Hendley, "In silico assessment of histotripsy-induced changes in catheter-directed thrombolytic delivery," *Front. Physiol.*, vol. 14, no. June, pp. 1–14, 2023, doi: 10.3389/fphys.2023.1225804.

[50] E. Vlaisavljevich *et al.*, "Non-Invasive Liver Ablation Using Histotripsy: Preclinical Safety Study in an In~Vivo Porcine Model," *Ultrasound Med. Biol.*, vol. 43, no. 6, pp. 1237–1251, Jun. 2017.

[51] B. D. de Senneville, C. Moonen, and M. Ries, "MRI-Guided HIFU Methods for the Ablation of Liver and Renal Cancers," in *Therapeutic Ultrasound*, J.-M. Escoffre and A. Bouakaz, Eds. Cham: Springer International Publishing, 2016, pp. 43–63.

[52] G. P. L. Thomas, T. D. Khokhlova, and V. A. Khokhlova, "Partial Respiratory Motion Compensation for Abdominal Extracorporeal Boiling Histotripsy Treatments with a Robotic Arm," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 68, no. 9, pp. 2861–2870, 2021, doi: 10.1109/TUFFC.2021.3075938.

[53] A. Denys, Y. Lachenal, R. Duran, M. Chollet-Rivier, and P. Bize, "Use of high-frequency jet ventilation for percutaneous tumor ablation," *Cardiovasc. Intervent. Radiol.*, vol. 37, no. 1, pp. 140–146, 2014, doi: 10.1007/s00270-013-0620-4.

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