IEEE TRANSACTIONS ON ULTRASONICS, FERROELECTRICS, AND FREQUENCY CONTROL, VOL. 70, NO. 9, SEPTEMBER 2023

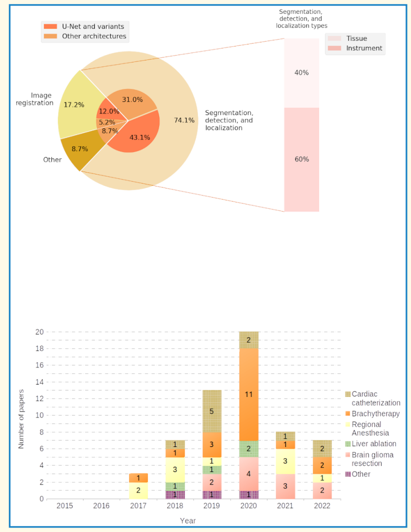
Author accepted manuscript version of the publication in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, volume 70, issue 9, September 2023. Published version: https://doi.org/10.1109/TUFFC.2023.3255843

The Big Bang of Deep Learning in Ultrasound-Guided Surgery: A Review

Nima Masoumi, Hassan Rivaz, *Senior Member, IEEE*, Ilker Hacihaliloglu, *Member, IEEE*, M. Omair Ahmad, *Life Fellow, IEEE*, Ingerid Reinertsen, and Yiming Xiao, *Senior Member, IEEE*

***Abstract*—Ultrasound (US) imaging is a paramount modality in many image-guided surgeries and percutaneous interventions, thanks to its high portability, temporal resolution, and cost-efficiency. However, due to its imaging principles, the US is often noisy and difficult to interpret. Appropriate image processing can greatly enhance the applicability of the imaging modality in clinical practice. Compared with the classic iterative optimization and machine learning (ML) approach, deep learning (DL) algorithms have shown great performance in terms of accuracy and efficiency for US processing. In this work, we conduct a comprehensive review on deep-learning algorithms in the applications of US-guided interventions, summarize the current trends, and suggest future directions on the topic.**

***Index Terms*—Deep learning (DL), intervention, percutaneous, surgical guidance, ultrasound (US).**



I. Introduction

Ultrasound (US) is a non-ionizing imaging modality that is commonly employed in the clinic, offering 2-D, 3-D, and 4-D data. Although US transducers are often operated in a free-hand manner by a physician or technician, to ensure image quality, semi-automatic or fully automatic image acquisitions are performed with the assistance of robotic arms in some applications [1]. While avoiding radiation risks, US scanners are portable and cost-effective as opposed to other staple imaging techniques, such as magnetic resonance imaging (MRI) and computed tomography (CT). In addition, US offers real-time anatomical and physiological information with great flexibility in applications, such as endoscopic, laparoscopic, transrectal, and transvaginal imaging. In addition to the most commonly seen B-mode contrast for structural imaging, US also provides additional contrasts, including Doppler US for flow imaging and US elastography computed from raw radio frequency (RF) scans to visualize the biophysical properties of tissues. These advantages of US imaging make it a favorable modality for image-guided interventions, where it is used for instrument and biological tissue (e.g., lesions) detection and tracking [2], [3].

Despite multiple benefits, US still faces several drawbacks primarily as a result of its inherent imaging principle. First, US scans are often noisy and prone to imaging artifacts such as reverberations, refraction, and shadowing, making recognition of anatomy and surgical tools difficult at times. Second, US usually has limited imaging depth, which can restrict the field of view for inspecting the pathological region. Finally, unlike modalities such as MRI and CT that have standardized planes, the unique image contrast and arbitrary and unfamiliar imaging planes make it challenging to interpret US images. So far, a great number of image processing techniques were proposed to tackle these aforementioned drawbacks, including denoising [4], structure or instrument detection [5], [6], segmentation [7], and image registration

Manuscript received 15 January 2023; accepted 6 March 2023. Date of publication 10 March 2023; date of current version 29 August 2023. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), in part by the Fonds de Recherche du Quebec Nature and Technologies (FRQNT), and in part by the Regroupment Strategique en Microelectronique du Quebec. *(Corresponding author: Hassan Rivaz.)*

Nima Masoumi, Hassan Rivaz, and M. Omair Ahmad are with the Department of Electrical Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: n_masoum@encs.concordia.ca; hrivaz@ece.concordia.ca; omair@ece.concordia.ca).

Ilker Hacihaliloglu is with the Department of Radiology, Department of Medicine, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: ilker.hacihaliloglu@ubc.ca.)

Ingerid Reinertsen is with the Department of Health Research, SINITEF Digital, 7465 Trondheim, Norway (e-mail: ingerid.reinertsen@sintef.no).

Yiming Xiao is with the Department of Computer Science and Software Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: yiming.xiao@concordia.ca).

This article has supplementary downloadable material available at https://doi.org/10.1109/TUFFC.2023.3255843, provided by the authors.

Digital Object Identifier 10.1109/TUFFC.2023.3255843

Highlights

- **We conduct a comprehensive review on deep-learning algorithms in the applications of ultrasound-guided interventions, summarize the current trends, and suggest future directions.**
- **Near 74% of reviewed methods perform segmentation, detection, and localization of medical instruments and target tissues, wherein U-Net and its variants were employed more than other models.**
- **With the ability to further reduce the demand for data annotation, unsupervised learning may hold an important role for future developments in interventional applications.**

[8], [9], [10], [11]. Traditionally, these techniques heavily rely on time-consuming iterative optimization methods or suboptimal hand-crafted features for classic machine learning (ML) algorithms. In comparison to conventional techniques, deep learning (DL)-based methods have shown excellent results in many US processing tasks by leveraging the computing power of modern graphics processing units (GPUs) [12], [13]. In addition, DL-based methods are faster at inference time, especially for large images [14]. With high requirement in accuracy, robustness, and efficiency, DL is well suited to facilitate US-guided interventions. To facilitate readers from diverse backgrounds, we have included a concise introduction to DL in Section S1 of the Supplementary Material.

To date, a number of literature reviews have been conducted on the topic of US-guided interventions. However, most of these previous reviews focus on the survey of clinical applicability of intra-operative US [15], [16], [17] or related acquisition techniques [18], [19], [20], [21]. With the great promise of DL techniques to enhance the value of intra-operative US, it is beneficial to provide a comprehensive review of the advancement of DL techniques in therapeutic interventional US. Based on the survey, we also identify the unmet clinical needs and suggest future research directions in the domain.

II. Literature Selection

We searched the literature using the Google Scholar database. The search was performed for publications from January 2015 to December 2022, the period that DL-based techniques gain popularity in medical imaging. The search criteria "US AND (Guided OR Surgery OR Intraoperative) OR (Convolution OR DL)" was utilized. The articles reviewed are on the technical development and validation of the algorithms, and review articles, case reports, and clinical reports are excluded from the search. The selected papers were carefully screened to ensure they were relevant to US-guided surgery and percutaneous interventions. US-guided diagnosis and biopsies were excluded from our search to focus our review on the therapeutic application of US imaging. The survey resulted in 58 papers. A breakdown of reviewed papers' numbers for each year is shown in Fig. 1. A breakdown of the reviewed DL methods in this study is illustrated in Fig. 2. To help the readers with their technical developments, we conducted a brief introduction to the common DL models in this survey in Section S2, a summary of the public datasets used in the reviewed papers in Section S3, and a list of reviewed papers' public codes with the web links in Section S4 of the Supplementary Material.

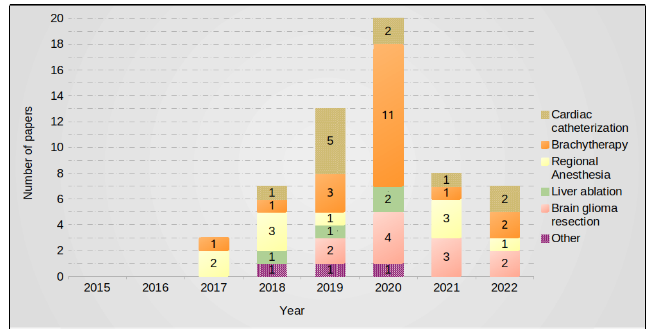


Fig. 1. Breakdown of reviewed papers' numbers for each year is presented. In total, 58 papers were studied. We did not find relevant publications in 2015 and 2016. The number of DL-based approaches in US-guided therapeutic interventions has grown from 2016 until 2020. The drop in publications in the year 2021 is likely due to the COVID-19 pandemic, which may have substantially impacted new data acquisition and research progress in the domain.

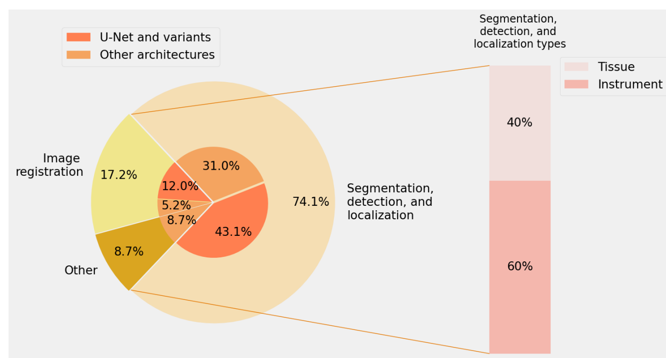


Fig. 2. Methods were classified into three categories: 1) segmentation, detection, and localization; 2) image registration; and 3) other methods. Since U-Net and its variants were employed dominantly more than the other models, we divide the utilized models into: 1) U-Net and variants and 2) other architectures. Most methods perform segmentation, detection, and localization of medical instruments and target tissues. These methods can be further broken down into tissue and instrument segmentation, detection, and localization. The other methods include the classification of tissues, motion detection, and so on.

III. Clinical Applications

The main clinical applications of the reviewed papers are US-guided cardiac catheterization, brachytherapy, regional anesthesia, liver ablation, and brain glioma resection. While most papers focus on one application, the others validated the proposed techniques in multiple. Since typically different

TABLE I

SUMMARY OF DL-BASED METHODS FOR HEART CATHETERIZATION IS PRESENTED. THE METHODS ARE MOSTLY FOCUSED ON CATHETER SEGMENTATION. THE EXAMINED DATASETS ARE ALL PRIVATE

Reference	Task	Proposed approach	Dataset	Key metric and performance
Yang *et al.* [22]	Catheter localization	A CNN with binary pre-selection of candidate voxels, and applied a Frangi vesselness filter [33] with adaptive thresholding	3D *ex-vivo* porcine heart US	Hausdorff distance of $1.64 \pm 1.82 voxels$
Yang *et al.* [23]	Instrument localization	A modified multi-scale U-Net [34] with a hybrid loss consisting of a contextual loss and a class-balanced focal loss	3D *ex-vivo* porcine heart US and 3D US of *in-vivo* human heart during TAVI operations	Dice score (%) of 69.6 ± 10.9 for *ex-vivo* and 65.8 ± 9.2 for *in-vivo* data
Yang *et al.* [24]	Catheter localization	A 3D U-Net [34] with a cross-entropy focal loss	3D *ex-vivo* porcine heart US	Skeleton error of $1.28mm$
Yang *et al.* [25]	Catheter detection	An early fusion CNN and a late fusion CNN [35] with a weighted cross-entropy loss	3D *ex-vivo* porcine heart US	Position error of $1.7voxels$
Yang *et al.* [26]	Instrument segmentation	Path-of-interest selection with fusion of a Pyramid-UNet [23] and a direction-fused U-Net which is based on a VGG16 encoder [36]	3D *ex-vivo* porcine heart US and 3D US of *in-vivo* human heart during TAVI operations	Dice score (%) of 70.5 ± 9.2 for *ex-vivo* and 66.5 ± 8.3 for *in-vivo* data
Yang *et al.* [27]	Instrument segmentation	Semi-supervised learning of a deep Q-network using a hybrid loss that consists of uncertainty and contextual constraints	3D *ex-vivo* porcine heart US and 3D US of *in-vivo* human heart during TAVI operations	Dice score (%) of 69.1 ± 7.3 for *ex-vivo* and 68.6 ± 7.9 for *in-vivo* data
Yang *et al.* [28]	Catheter segmentation	Weakly-supervised learning using a ResNet10 encoder [37] with the class attention maps-guided [38] pseudo-label generation	3D *ex-vivo* porcine heart frustum US	Dice score (%) of 65.4 ± 9.7
Yang *et al.* [31]	Catheter segmentation and localization	A direction-fused U-Net which is based on a VGG16 encoder [36]	3D *ex-vivo* porcine heart US	Dice score (%) of 67.7 ± 12.0
Min *et al.* [32]	Catheter segmentation	A VGG encoder [36] with pre-selection of candidate voxels, and applied a Frangi vesselness filter [33]	3D *ex-vivo* porcine heart frustum US	Dice score (%) of 67.3 ± 14

surgical procedures have different needs, the review of the developed techniques is conducted with respect to their clinical applications.

A. US-Guided Cardiac Catheterization

Catheterization is common in various cardiac interventions, such as angioplasty and heart valve surgery. The catheter has a narrow tubular shape inserted into the patient's artery. The intraoperative X-ray is commonly acquired to localize the catheter. X-ray imaging has risks for interventionalists and patients due to its ionizing radiation. Given this fact, a safer choice, US-guided catheterization, is gaining popularity over intraoperative X-rays. However, locating the catheter in US images, especially near the heart chamber, is challenging, and in the clinic, fast uptake is required. Robust image processing algorithms can automatically detect and localize the catheter in US images. Furthermore, they can also perform voxel/pixel-wise segmentation of the catheter with submillimeter precision. Yang et al. [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], and [32] in several studies, showed that DL approaches could help the localization and detection of the catheter in US images. They proposed methods to segment pixels/voxels into catheter and non-catheter classes. The methods were validated in several applications, such as transcatheter aortic valve implantation (TAVI). The methods are summarized in Table I, and they are primarily validated using private 3-D ex vivo animal and in vivo human datasets. In terms of instrument segmentation, these methods achieved Dice scores up to 70%.

B. US-Guided Brachytherapy

Brachytherapy is a procedure for treating certain kinds of cancers. In this procedure, small radioactive seeds are placed into the target region of the patient's body using needles or a catheter. The radiation dose of seeds in brachytherapy should be well-localized to the pathological region and spares the adjacent healthy tissues. Therefore, intra-operative guidance, especially with US has gained the attention of radiation therapists. For prostate brachytherapy, transrectal US (TRUS) is commonly used to guide multiple medical instruments to the targeted region for the correct placement of seeds. Multi-needle localization, detection, and segmentation in US images can help accurate insertion of radioactive seeds and potentially improve the treatment efficacy and safety. Ideally, automatic algorithms that perform these tasks should operate in real-time and be robust against image noise and signal distortion in real clinical applications. Zhang et al. [39] and [40], in two different studies, proposed multi-needle localization methods using an attention U-Net [34] and a region-based convolutional neural network (R-CNN). They validated their method on 3-D TRUS of patients who underwent high-dose-rate (HDR) brachytherapy. A CNN model was developed by Andersén et al. [41] to digitize needles in 3-D TRUS of prostate HDR brachytherapy patients. Wang et al. [42] proposed a U-Net and an additional VGG16-based deep network to segment brachytherapy needles in 3-D volumes reconstructed from 2-D TRUS slices. Liu et al. [43] trained and tested a U-Net model to localize catheter in 3-D reconstructed TRUS images taken during several prostate HDR brachytherapies.

Intraoperative prostate segmentation can facilitate the treatment target identification in consideration of the patient motion, thus improving the efficiency, safety, and therapeutic outcomes. Girum et al. [44] and [45] proposed DL approaches using a U-Net and a generative CNN to segment the prostate in 3-D reconstructed volumes from 2-D TRUS slices. Orlando et al. [46] proposed a DL method using a modified U-Net for prostate segmentation on clinically diverse 3-D TRUS images. Later, they developed two DL methods using a modified U-Net and a U-Net++ [47], [48], which were trained on 2-D TRUS slices [49]. Nevertheless, the methods were tested on 3-D TRUS volumetric images. Lei et al. [50] proposed DS-V-Net which is a prostate segmentation method using multidirectional V-Net [51]. The popular DS-V-Net achieved the Dice score (%) of 91.9 $\pm$ 2.8 in clinical data. A prostate target volume delineation method using residual neural networks for low-dose-rate brachytherapy was developed by Anas et al. [52]. The method was validated on 2-D TRUS slices using manual segmentation as ground truths. Karimi et al. [53] and [54] proposed a novel CNN architecture for prostate segmentation in 2-D TRUS images. He et al. et al. [55] proposed a deep-attentional GAN-based method to improve the resolution of 3-D TRUS images. Golshan et al. [56] proposed a modified LeNet architecture [57] for radioactive seeds segmentation in 3-D TRUS images. This will help confirm the location of implantation and facilitate the procedure that removes these seeds after the treatment period.

Pre-operative MRIs are often acquired for surgical planning in prostate brachytherapy. MRIs generally have sharper images and show better details of the target area than US. Image registration of intraoperative TRUS with the MRI can help guide the interpretation of the US scans. Chen et al. [58] proposed a DL approach using V-Net and U-Net architectures to segment and register the prostate in MR and TRUS. Zeng et al. [59] performed 3-D non-rigid registration of MR-TRUS using convolutional and recurrent neural networks.

Brachytherapy is not confined to prostate cancer treatment. Rodgers et al. [60] proposed a DL-based method for needle localization in 3-D transvaginal US images of interstitial gynecologic HDR brachytherapy. Sun et al. [61] generated pseudo-CT images from intraoperative US images of cervical cancer patients for brachytherapy. The DL methods in this section are summarized in Table II. The Dice score, followed by the shaft and needle tip localization errors, are the key metrics for quantitative performance assessment. In general, the algorithms achieved sub-millimeter accuracy in shaft and needle tip localization.

C. US-Guided Regional Anesthesia

Needle-based regional anesthesia is conventionally used in operating rooms. It usually requires an experienced expert to deliver the anesthetic injection. US-guided regional anesthesia can help the anesthesiologist with the procedure. Detection and localization of the injection needle shaft and tip can be challenging. In 2-D US scans, needle tips are occasionally out-of-plane or difficult to spot. Processing raw US RF data or 3-D reconstructed scans is helpful for accurate and reliable needle identification. DL approaches can help with needle localization in US images [67]. Mwikirize et al. [68], [69], and [70] developed CNNs in three studies to localize the needle tip in real-time 2-D US images. Gao et al. [71] proposed a needle segmentation method using a U-Net architecture. Pourtaherian et al. [72] proposed a needle tip detection method using orthogonal-plane CNNs. They validated their method on ex vivo 3-D US images of chicken breast. Later, they developed a method for the localization of needle tips with sub-millimeter accuracy using dilated CNNs [73]. Finally, Maneas et al. [80] modified an established residual neural network to improve the axial and lateral resolution of tracked US images for needle localization. They trained their model on synthetic data, and the model was validated on a fetal sheep heart in vivo data.

Nerve segmentation in US scans for US-guided regional anesthesia can facilitate the practitioners with the procedure. Automatic non-learning methods using Kalman filters could rapidly perform nerve and artery segmentations [74]. Generally, these methods are computationally expensive and require intensive hyperparameter tuning but recent works proposed DL-based techniques to address the drawbacks of classic Kalman filtering [75], [76]. Smistad et al. [77] proposed a technique using U-Net for musculocutaneous, median, ulnar, and radial nerve segmentation during axillary nerve block procedures. Baby and Jereesh [78] developed a U-Net model to delineate the brachial plexus in 2-D US images. A conditioned U-Net model (www.kaggle.com/harolddiaz1018/unet-cond) was trained by Díaz-Vargas et al. [79] to segment ulnar, median, femoral, and sciatic nerves in 2-D US slices. The DL methods in this section are summarized in Table III. The shaft and needle tip localization errors are the key metrics for quantitative performance assessment.

D. US-Guided Liver Ablation

Image-guided microwave ablation (MWA) is a promising therapeutic percutaneous intervention that provides a high intralesional temperature. Real-time US imaging techniques can visualize the target for accurate lesion MWA and complete tumor eradication. However, the ablation region margin is not easily detectable in US images. While ablation region delineation is feasible using techniques such as US elastography [85], we focus our review on DL techniques. Unsupervised classification of target region tissues by leveraging an ML/DL-based method is a candidate approach. Zhang et al. [86] utilized the raw US RF data and trained a CNN network to delineate the ablation region in ex-vivo data of the porcine liver. Wang et al. [87] proposed a CNN-based method for ablation region detection and monitoring MWA. They performed image registration of US RF data and optical images to boost the accuracy of their method in terms of receiver operating characteristic curves. Kondo et al. [88] proposed an out-of-plane motion detection system using CNNs to track liver tumor movement in ablation therapies.

Ablation needle detection and visualization can help interventionalists during the MWA procedure. Arif et al. [89] proposed a real-time bi-planar needle detection and visualization for liver 3-D US images. Their method utilizes a U-Net architecture and specific post-processing to perform the needle detection. They execute the registration of images in different time frames acquired from liver phantom and ten patients. The DL methods in this section are summarized in Table IV. Dice score and mean absolute error are the key metrics for quantitative performance assessment.

E. US-Guided Brain Glioma Resection

US scanners' portability and cost-effectiveness of US imaging contributed to the growing popularity of intraoperative US

TABLE II
SUMMARY OF DL-BASED METHODS FOR US-GUIDED BRACHYTHERAPY IS PRESENTED. THE METHODS ARE MOSTLY FOCUSED ON TARGET AND INSTRUMENT SEGMENTATION. PUBLIC DATASETS ARE MARKED WITH "*"

Reference	Task	Proposed Approach	Dataset	Key metrics and Performance
Zhang *et al.* [39]	Multi-needle localization	A deep supervised attention U-Net with a weighted total variation regularization	3D *in-vivo* TRUS of prostate HDR brachytherapy	Shaft localization error of $0.29 \pm 0.23mm$ and needle tip localization error of $0.44 \pm 0.83mm$
Zhang *et al.* [40]	Multi-needle localization	A reformulated large-margin mask R-CNN [62] that is combined with a density-based spatial clustering [63]	3D *in-vivo* TRUS of prostate HDR brachytherapy	Shaft localization error of $0.09 \pm 0.04mm$ and needle tip localization error of $0.33 \pm 0.36mm$
Andersen *et al.* [41]	Digitization of prostate brachytherapy needles	A 3D U-Net architecture [34] with a Dice loss	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	Root-mean-square deviation (RMSD) median (interquartile range) of $0.55 \quad (0.35 - 0.86)$
Wang *et al.* [42]	Needle segmentation	A U-Net followed by a VGG16 network with a categorical cross-entropy loss	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	Resolution of needle trajectories of $0.66mm$ and $0.31mm$ in x and y direction respectively
Liu *et al.* [43]	Catheter localization	A U-Net architecture with a focal Tversky loss function [64]	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	80% within $2mm$ catheter reconstructions
Girum *et al.* [44]	Prostate clinical target-volume boundary detection	A modified U-Net with an integrated regression model using global average pooling	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	Dice score (%) of 88.0 ± 2.0
Girum *et al.* [45]	Prostate clinical target-volume segmentation	A CNN for prior-knowledge generator and a CNN for the segmentation	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy, 3D *in-vivo* postoperative CT scans of prostate HDR brachytherapy, and *2D *in-vivo* echocardiography images	Dice score (%) of 96.9 ± 0.9, 95.4 ± 0.9, and 96.3 ± 1.3 for TRUS, CT, and 2D echocardiography images respectively
Orlando *et al.* [46]	Prostate segmentation	A 2D modified U-Net with a Dice loss	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy and biopsy	A median (first quartile - third quartile) absolute Dice score (%) of $94.1 \quad (92.6 - 94.9)$
Orlando *et al.* [49]	Prostate segmentation	Trained U-Net and U-Net++ [47], [48] architectures separately using 2D slices	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy and biopsy	A median (first quartile - third quartile) absolute Dice score (%) of $94.1 \quad (92.6 - 94.9)$ and $94.0 \quad (92.2 - 95.1)$ for U-Net and U-Net++ respectively
Lei *et al.* [50]	Prostate segmentation	A multidirectional deeply supervised V-Net [51] with a hybrid loss that consists of a binary cross-entropy loss and a batch-based Dice loss	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate	Dice score (%) of 91.9 ± 2.8
Anas *et al.* [52]	Clinical target-volume delineation	CNNs based on ResNets [37] and dilated convolution at deeper layers	2D *in-vivo* TRUS of prostate brachytherapy patients	Dice score (%) of 93.67 ± 3.71
Karimi *et al.* [53], [54]	Clinical target-volume segmentation	Sparse subspace clustering [65] of features learned with a convolutional auto-encoder and a modified U-Net architecture	2D *in-vivo* TRUS of prostate brachytherapy patients	Dice score (%) of 93.9 ± 3.5
Xiuxiu *et al.* [55]	Improving US image resolution	Integrating a deeply supervised attention model into a generative adversarial network-based	3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate	Mean absolute error of 6.5 ± 0.5
Golshan *et al.* [56]	Brachytherapy seeds detection	A LeNet [57] architecture with a cross-entropy loss	3D *in-vivo* volumes reconstructed from 2D original TRUS RF data of prostate brachytherapy patients	Precision, recall, and F1-score (%) of 78.0 ± 8.0, 64.0 ± 10.0, and 70.0 ± 8.0 respectively
Chen *et al.* [58]	MR to TRUS image registration and prostate segmentation	Segmentation-based registration using two weakly-supervised 3D V-Nets [51] for segmentations and a 3D U-Net for the registrations	3D *in-vivo* T2w MRI and 3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	Dice score (%) of 97.0 ± 0.0 and 87.0 ± 5.0 for segmented mask and manual contours respectively
Zeng *et al.* [59]	MR to TRUS prostate registration	A modified U-Net [34] and a bidirectional convolutinoal LSTM with a hybrid loss that consists of a bending energy loss and a Dice loss	3D *in-vivo* T2w MRI and 3D *in-vivo* TRUS volumes reconstructed from 2D slices of prostate HDR brachytherapy	Dice score (%) 90.0 ± 4.0
Rodgers *et al.* [60]	Needle localization	A 2D U-Net [83] for 2D data and randomized 3D Hough transforms [66] for 3D data	3D *in-vivo* transvaginal US (TVUS) volumes reconstructed from 2D slices of interstitial gynecological HDR brachytherapy, 2D US slices of phantom brachytherapy, and 2D US slices of ablation therapy	Median position difference (first quartile - third quartile) of $0.27 \quad (0.20 - 0.68)mm$ and $0.79 \quad (0.62 - 0.93)mm$ for 2D and 3D TVUS respectively
Sun *et al.* [61]	Pseudo-CT image synthesis from US	A part of VGG19 [36] network and a hybrid loss that consists of a content loss, a style loss, and a Dice loss	3D *in-vivo* CT scans and 3D *in-vivo* US volumes of cervical cancer patients, additional 3D CT scans of cervical cancer patients, and 3D US phantom data	T-test of structural similarity index between the ground-truth and synthesized CT with $t = 3.22$ and $t = 2.86$ for background and foreground regions respectively

acquisition. Spatially tracked US probes can be calibrated and synced with a neuronavigation system in operating rooms to allow the overlay of real-time US scans with pre-operative surgical plans. Practitioners may execute image registration between preoperative images and intraoperative US to update the surgical plan. For instance, in brain glioma surgery, intra-operative US images can be registered to the preoperative MRIs (or intraoperative US images at different time points). Because after surgeons open the dura, the brain tissue can deform up to 18 mm due to several causes, including gravity, cerebrospinal fluid loss, drug administration, retraction, resection, and so on [90], [92]. This phenomenon is commonly called brain shift. Brain shift can make the preoperative planning invalid. Therefore, fast registration of preoperative and intraoperative data is crucial. Public datasets, such as the brain images of tumors for evaluation database (BITE) [90] and retrospective evaluation of cerebral tumors (RESECT) [92] databases have greatly facilitated the development of methods for brain-shift correction, including the DL approaches. In the CuRIOUS2018 Challenge held in conjunction with MICCAI

TABLE III

SUMMARY OF DL-BASED METHODS FOR US-GUIDED REGIONAL ANESTHESIA IS PRESENTED. ANESTHESIA NEEDLE TIP LOCALIZATION IS THE FOCUS OF THE MAJORITY OF WORKS. PUBLIC DATASETS ARE MARKED WITH "*"

Reference	Task	Proposed Approach	Dataset	Key metrics and Performance
Mwikirize *et al.* [68]	Real-time needle detection	A region-based CNN [81] and a region-proposal CNN	2D *ex-vivo* US bovine and porcine tissues, and 2D US of bovine/porcine tissues overlaid on lumbosacral spine phantom	Shaft localization error of $0.82° \pm 0.4°$, and needle tip localization error of $0.23 \pm 0.05mm$
Mwikirize *et al.* [69]	Real-time needle tip localization	Needle enhancement followed by a CNN for needle tip classification, and a CNN regression network	2D *ex-vivo* US of bovine, porcine, and chicken tissues overlaid on lumbosacral spine phantom	Needle tip localization error of $0.55 \pm 0.07mm$
Mwikirize *et al.* [70]	Needle tip localization	A novel network that consists of convolutional layers and recurrent layers (CNN-LSTM) with a Mean Squared Error (MSE) loss	2D *ex-vivo* US of bovine, porcine, and chicken tissues overlaid on lumbosacral spine phantom	Needle tip localization error of $0.52 \pm 0.06mm$
Gao *et al.* [71]	Needle localization and enhancement	Beam steering followed by a modified U-Net for segmentation, and a needle fusion algorithm	2D *ex-vivo* US of bovine, porcine, and chicken tissues	Needle tip localization error of $0.29 \pm 0.02mm$
Pourtaherian *et al.* [72]	Needle detection	Two CNNs with shared and independent convolutional filters respectively using a categorical cross-entropy cost	3D *ex-vivo* US of a chicken breast	Precision 83% at 76% recall
Pourtaherian *et al.* [73]	Needle localization	CNNs with dilated convolutions and spatial pyramid pooling features	3D *ex-vivo* US of a porcine leg	Qualitative assessment
Esmistad *et al.* [77]	Nerve segmentation	A modified U-Net	2D *in-vivo* US of volunteers and patients undergoing axillary nerve block procedures	Precision of 88%, 63% 79%, 67%, and 44%, and recall of 0.81, 0.56, 0.71, 0.62, and 0.37 for blood vessel, musculocutaneous nerve, median nerve, ulnar nerve, and radial nerve respectively
Baby *et al.* [78]	Nerve segmentation	A modified U-Net	*2D *in-vivo* US of patients' brachial plexus	Dice score 71%
Diaz-Vargas *et al.* [79]	Peripheral nerve segmentation	A conditioned U-Net with a Dice loss	2D *in-vivo* US of patients' ulnar, median, femoral, and sciatic nerves	Dice score (%) of 70.0 ± 27.0
Maneas *et al.* [80]	Instrumented ultrasonic tracking	ResNet architecture [84] with a L1-loss	2D synthetic US RF data, and 2D *in-vivo* US of fetal sheep heart	Root-mean-square error of 0.15 ± 0.03 for the synthetic data, and signal energy concentration ration of 99.9% for the *in-vivo* data

TABLE IV

SUMMARY OF DL-BASED METHODS FOR US-GUIDED LIVER ABLATION IS PRESENTED. THE EXAMINED DATASETS ARE ALL PRIVATE

Reference	Task	Proposed Approach	Dataset	Key metrics and Performance
Zhang *et al.* [86]	Thermal lesion detection	Matched pathology images to US RF data followed by training a CNN with two paths	2D *ex-vivo* US B-mode and RF data liver tissues	Dice score 86.88%
Wang *et al.* [87]	Thermal lesion detection	Image registration of RF data and optical images followed by training a CNN	2D *ex-vivo* US B-mode and RF data, and optical images of the porcine liver tissues	Receiver operating characteristic curve of 0.87
Kondo *et al.* [88]	Tumour motion detection	A VGG16 [36] followed by a U-Net architecture with a hybrid loss	2D US of liver phantom	Mean absolute error of $0.342mm/frame$
Arif *et al.* [89]	Needle detection	Image registration of needles in different time points and needle segmentation using a compressed V-Net [51]	3D *in-vivo* US of liver biopsy patients, and 3D US of puncturing phantoms	Mean absolute error of $1.00mm$ and $2.0°$ for needle position and orientation respectively

2018, the participating teams were asked to register preoperative MRI to intraoperative US images of the RESECT dataset. The challenge results and participating teams' methods are summarized and compared in [91] with most methods using traditional approaches. Canilini et al. [93] proposed a DL method using a CNN to segment sulci and falx cerebri in US images. Then, they used the segmentation masks to register intraoperative, preoperative, and postoperative US images. The method was tested on BITE and RESECT datasets. Given the fact that these datasets provide manual homologous landmarks, Canilini et al. [94] calculated mean target registration error (mTRE) for the quantitative validation of their method. Later, they trained a U-Net architecture to generate segmentation masks of resection cavities. They registered the US volumes using these masks.

Zeineldin et al. [95], [96], and [97] proposed DL-based methods with U-Net architectures in different studies to register preoperative MRI to intraoperative US images. They employed MSE Loss for their model training in [95]. Later, they used MSE loss and NCC loss in a comparison study in [96] and NCC loss in [97]. Pirhadi et al. [98] employed a Siamese neural network [99] to perform landmark-based registration of pre-resection intraoperative US to post-resection intraoperative US scans.

Finding the precise boundaries of the tumor and its segmentation can assist surgeons to optimize the resection boundary. Zeineldin et al. [100] employed U-Net and TransUNet networks [101] to segment brain tumors in US images. Carton et al. [102] proposed a DL-based method with a 3-D U-Net architecture to segment the brain tumors of RESECT dataset intraoperative US images. In addition to lesion segmentation, classification of the lesion into different glioma grades or solitary brain metastases can be substantial because the surgical procedures vary for each case. Cepeda et al. [103] proposed a DL approach to analyze the candidate lesions in patients who underwent craniotomy. They used B-mode and strain elastography images to correctly classify the lesions as glioblastoma or solitary brain metastases. The DL methods in this section are summarized in Table V.

F. Other Clinical Applications

Sections III-E reviewed the DL approaches in widely studied clinical applications. This section reviews the clinical

TABLE V

SUMMARY OF DL-BASED METHODS FOR US-GUIDED BRAIN GLIOMA RESECTION IS PRESENTED. MOST METHODS PERFORM IMAGE REGISTRATION FOR BRAIN SHIFT CORRECTION IN BITE [90] AND RESECT [92] DATASETS. PUBLIC DATASETS ARE MARKED WITH "*"

Reference	Task	Proposed Approach	Dataset	Key metrics and Performance
Canilini *et al.* [93]	Segmentation and registration of US volumes	Segmentation by a modified U-Net [34] and registration of generated masks	*3D *in-vivo* US volumes reconstructed from 2D slices of RESECT [92] and BITE [90] datasets	mTRE of 2.05 $\pm$ 1.12mm for RESECT and 2.48 $\pm$ 2.67mm for BITE dataset
Canilini *et al.* [94]	Resection cavity segmentation and registration of US volumes	Segmentation by a modified U-Net [34] and registration of generated masks	*3D *in-vivo* US volumes reconstructed from 2D slices of RESECT [92] and BITE [90] datasets	mTRE of 1.21 $\pm$ 0.66mm for volumes before and after resection of RESECT, 1.22 $\pm$ 1.20mm for volumes before and during resection of RESECT, and 2.38 $\pm$ 2.78mm for BITE dataset
Zeineldin *et al.* [95]	MR to US registration	A U-Net architecture with a MSE loss	*3D *in-vivo* US volumes reconstructed from 2D slices and 3D T2-FLAIR MRI of RESECT [92]	Mean squared error of 85
Zeineldin *et al.* [96]	MR to US registration	Two U-Net architecture with MSE and NCC losses respectively	*3D *in-vivo* US volumes reconstructed from 2D slices and 3D T2-FLAIR MR of RESECT [92] and BITE datasets	mTRE of 0.84 $\pm$ 0.16mm for RESECT and 1.47 $\pm$ 0.61mm for BITE dataset
Zeineldin *et al.* [97]	MR to US registration	A U-Net architecture with a NCC loss	*3D *in-vivo* US volumes reconstructed from 2D slices and 3D T2-FLAIR MRI of RESECT [92] and BITE datasets	mTRE of 0.99 $\pm$ 0.22mm for RESECT and 1.68 $\pm$ 0.65mm for BITE dataset
Pirhadi *et al.* [98]	Landmark-based US volumes registration	A Siamese network [99] for detecting landmarks with a 2.5D approach [104]	*3D *in-vivo* US volumes reconstructed from 2D slices of RESECT [92] and BITE [90] datasets	mTRE of 1.22 $\pm$ 0.46mm for volumes before and after resection of RESECT, 1.11 $\pm$ 0.43mm for volumes before and during resection of RESECT, and 1.76 $\pm$ 1.48mm for BITE dataset
Zeineldin *et al.* [100]	Brain tumour segmentation	U-Net [83] and TransUNet [101] architectures	*3D *in-vivo* US volumes reconstructed from 2D slices of RESECT [92] dataset	Dice scores (%) of 93.50 and 93.70 for U-Net and TransUNet respectively
Carton *et al.* [102]	Brain tumour segmentation	Three U-Net networks with Dice losses	*3D *in-vivo* US volumes reconstructed from 2D slices of RESECT [92] dataset	Median Dice score (%) of 72.00
Capeda *et al.* [103]	Glioblastoma and solitary brain metastases classification	Employed Inception V3 network from Orange software version 3.26 (University of Ljubljana, Slovenia)	2D *in-vivo* US images of supratentorial tumour patients who underwent craniotomy	Classification accuracy values of 0.79 to 0.94 for B-mode US and 0.84 to 0.97 for elastography data

applications with a few DL-based approaches. Lee et al. [105] proposed a DL method to classify liver fibrosis. They utilized the data for patients who underwent liver resection surgery. Gillies et al. [106] employed a U-Net architecture with a Dice loss to detect general interventional tools in 2-D US images. They utilized the datasets of prostate and interstitial gynecologic brachytherapy, liver, and kidney ablations. Wang et al. [107] proposed a deep attentive method for prostate segmentation. Their notable approach achieved the Dice score (%) of 90.0 $\pm$ 3.0 in the clinical target volume. Hu et al. [108] developed an adversarial deformation regularization method for preoperative and procedural TRUS image registration. However, the developed methods of Wang et al. [107] and Hu et al. [108] have not been designed for a focused application, and they can be used for prostate brachytherapy or prostatectomy.

IV. DISCUSSION AND FUTURE DIRECTIONS

Based on the literature included in the review, DL techniques have shown great promise to enhance the value of intra-operative US in surgical interventions. In most of the reviewed papers, the proposed DL methods were compared with traditional methods, where they showed that their techniques could significantly ($p < 0.05$) outperform the traditional ones in the execution time and the evaluation metrics. While segmentation, detection, and localization are the main techniques under development, these also need to be adapted to the application-specific needs and from the current state-of-the art, we identified a few unmet clinical needs that could be addressed by DL methods in the future. In the literature on brachytherapy, most efforts in DL techniques were dedicated to the prostate treatment, even though US-guided brachytherapy was also practiced for lung cancer, breast cancer, anal cancer, and abdominal wall metastases. Similarly, DL approaches in US-guided ablation are primarily focused on the liver while kidney and prostate ablation therapies still have the potentials for further technical development. In US-guided tumor resection procedures, similar DL methods can be further adapted for lumpectomy, prostatectomy, tongue cancer resection, laparotomy, pancreatic cancer resection, and bladder cancer resection. Finally, although, US was investigated as an intraoperative imaging tool in orthopedic surgery procedures, and complete system with extensive evaluation studies is still lacking. Currently, most focus in this domain has been given to developing accurate, robust, and fast bone segmentation [112], [113]. We believe efforts could be directed to propose and evaluate US bone registration approaches [114]. For some domain applications, such as cardiac catheterization, we found that the relevant works were mainly from a handful of labs. This may be due to the availability of clinical resources and collaboration, and it will be beneficial to have more confirmation studies from other research groups in the future. 3-D US volume reconstruction is critical for interventional guidance in many clinical applications, such as brain tumor resection [90], [92]. Leblanc et al. [109] proposed a US reconstruction technique for peripheral artery imaging. Luo et al. [110] leveraged a self-supervised strategy to reconstruct freehand 3-D US. Guo et al. [111] developed a learning model utilizing self-attention to reconstruct 3-D US volumes without tracking. However, most existing techniques use biopsy and diagnostic data to develop the algorithms due to their availability, but they can still be well applied in surgical applications.

Despite the excellent performance, DL techniques, including those reviewed in this article still have several drawbacks. First, most algorithms still require large well-annotated data to achieve good performance. This issue can be mitigated by adopting self-supervised and semi-supervised learning to learn feature representations by exploiting unlabeled or partially labeled data. Second, due to coarse and difficult-to-interpret image features as a result of US's imaging principle, accurate anatomical segmentation is often challenging. DL-based super-resolution and denoising techniques may help enhance the clarity of image features to mitigate the issue. Third, the trained networks often have limited adaptability to new domains (e.g., images from different scanner types or setting). Finally, most existing algorithms still lack transparency to help verify the quality of the outcomes. Currently, the lack of large-scale well annotated databases, especially the public repositories poses a bottleneck in algorithm development and fair performance benchmarking, and this also partially contributed to the various under-explored clinical applications as mentioned earlier, besides their application-specific challenges. In interventional applications, well-annotated data are often more difficult to obtain, especially with US. Currently to address the issue, weakly-supervised learning strategies in the reviewed papers have achieved impressive performance [27], [45], [59]. By leveraging categorical or coarse image annotations. With the ability to further reduce the demand for data annotation, unsupervised learning may hold an important role in future developments in interventional applications, but a more in-depth investigation is still required. In addition, data augmentation, including simulated US, can help overcome the scarcity of samples. However, the current techniques often fail to provide realistic results. Compared with MRI and CT scans, clear structural delineations in US is more difficult due to the nature of the imaging principles, and often co-registered biopsy, MRI, and CT data may be required when it comes to confirmation of pathological tissue segmentation. As direct contact is needed, for endoscopic applications, image acquisition also demands elaborate setup using surgical navigation systems or robotic assistance. These further complicate the construction of relevant datasets besides the privacy concerns commonly associated with medical data sharing.

In current literature, convolutional neural networks, especially different variants of U-Net architecture [83] have dominated the reviewed methods. In many applications, to overcome the limited data, CNNs previously trained with other imaging modalties (e.g., natural images) were adapted to the application domain with transfer learning [115]. However, partially due to the lack of public data, general-purpose DL algorithms that are more tolerant to scanner types and clinical applications still face major challenges. A few initiatives in MRI and CT DL registration and segmentation, such as the Learn2Reg MICCAI challenge [116] and the medical segmentation decathlon challenges [117] have attempted to help development these types of algorithms, but there is still a lack of similar endeavors in US. Accessibility to implementations facilitates transferring various architectures to new problems. As many learning-based approaches are highly data-dependent and application-tailored, efforts in the reproducibility of the published algorithms from the research community are still required to ensure the value of the technology in real practice. Several DL architectures are proposed in the reviewed literature. Optimal model selection can largely depend on various factors, including the suitability of data types (e.g., static versus temporal), data dimensions (e.g., 2-D versus 3-D), types of the target task (e.g., segmentation, registration, and so on), and requirement of portability (i.e., running on a mobile device, desktop computers, or cloud service). Besides decisions by human experts, automated DL model search has also attracted the attention of the research community [119]. However, automatic search strategies are still not widely adopted. The more recent vision transformers (ViT) have shown better performance in learning long-range spatial dependencies than CNNs, which require a more elaborate architecture design to model the spatial context of the image [118]. Adoption of ViT and its variants may further improve the accuracy of existing and future DL methods for intra-operative US.

Interpretability and trustworthiness of DL algorithms are crucial for the widespread adoption of the end products to the clinic. Conventional algorithms often have a goal-driven black-box design, and in this case, without careful verification, faulty automatic outputs can cause harms to the patients. The latest trend in explainable AI (XAI) intends to improve algorithm transparency through techniques, including spatial attention/activation visualization [120], [121], uncertainty estimation, and multi-task learning [122]. For various surgical applications, XAI methodologies can potentially further detect and explain problematic results from DL-based iUS processing that offer real-time feedback to improve the robustness and reliability of the algorithms, and thus the safety and efficiency of the surgery.

V. Conclusion

This review paper studied 58 DL-based approaches for US-guided heart catheterization, brachytherapy, regional anesthesia, liver ablation, and glioma resection. Near 74% of reviewed methods perform segmentation, detection, and localization of medical instruments and target tissues. Possible research directions for DL-based approaches were discussed in Section IV.

References

[1] J.-A. Long et al., "Development of a novel robot for transperineal needle based interventions: Focal therapy, brachytherapy and prostate biopsies," *J. Urology*, vol. 188, no. 4, pp. 1369–1374, Oct. 2012.

[2] A. Pourtaherian et al., "Medical instrument detection in 3-dimensional ultrasound data volumes," *IEEE Trans. Med. Imag.*, vol. 36, no. 8, pp. 1664–1675, Aug. 2017.

[3] S. G. Yuen, (2008), "3D ultrasound-guided motion compensation system for beating heart mitral valve repair," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2008*, vol. 5241, D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds. Berlin, Germany: Springer, 2008, pp. 711–719.

[4] F. P. X. de Fontes, G. A. Barroso, P. Coupé, and P. Hellier, "Real time ultrasound image denoising," *J. Real-Time Image Process.*, vol. 6, no. 1, pp. 15–22, Mar. 2011.

[5] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, Jan. 2010.

[6] W. Y. Chan, J. Qin, Y. P. Chui, and P. A. Heng, "A serious game for learning ultrasound-guided needle placement skills," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1032–1042, Nov. 2012.

[7] A. Belaid et al., "Phase-based level set segmentation of ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 138–147, Jan. 2011.

[8] H. Rivaz, S. J.-S. Chen, and D. L. Collins, "Automatic deformable mr-ultrasound registration for image-guided neurosurgery," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, pp. 366–380, Feb. 2015.

[9] H. Zhou and H. Rivaz, "Registration of pre- and postresection ultrasound volumes with noncorresponding regions in neurosurgery," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 5, pp. 1240–1249, Sep. 2016.

[10] N. Masoumi, Y. Xiao, and H. Rivaz, "ARENA: Inter-modality affine registration using evolutionary strategy," in *Proc. IJCARS*, vol. 14, 2019, 441–450, doi: 10.1007/s11548-018-1897-1.

[11] N. Masoumi, C. J. Belasso, M. O. Ahmad, H. Benali, Y. Xiao, and H. Rivaz, "Multimodal 3D ultrasound and CT in image-guided spinal surgery: Public database and new registration algorithms," in *Proc. IJCARS*, vol. 16, 2021, pp. 555–565, doi: 10.1007/s11548-021-02323-2.

[12] X. Chen, A. Diaz-Pinto, N. Ravikumar, and A. Frangi, "Deep learning in medical image registration," *Prog. Biomed. Eng.*, vol. 3, Dec. 2020, Art. no. 012003.

[13] A. K. Z. Tehrani, I. M. Rosado-Mendez, and H. Rivaz, "Robust scatterer number density segmentation of ultrasound images," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 4, pp. 1169–1180, Apr. 2022.

[14] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019.

[15] L. Dixon, A. Lim, M. Grech-Sollars, D. Nandi, and S. Camp, "Intraoperative ultrasound in brain tumor surgery: A review and implementation guide," *Neurosurgical Rev.*, vol. 45, no. 4, pp. 2503–2515, Aug. 2022.

[16] C. O. C. O. Danilo and S. Leanza, "Routine intraoperative ultrasound for the detection of liver metastases during resection of primary colorectal cancer—A systematic review," *Maedica*, vol. 15, no. 2, p. 250, 2020.

[17] B. Jiao, H. Chen, M. Chen, P. Lu, J. Liu, and C. Chen, "Opioid-sparing effects of ultrasound-guided erector spinae plane block for adult patients undergoing surgery: A systematic review and meta-analysis," *Pain Pract.*, vol. 22, no. 3, pp. 391–404, 2022.

[18] R. Kessner, D. A. Nakamoto, V. Kondray, S. Partovi, Y. Ahmed, and N. Azar, "Contrast-enhanced ultrasound guidance for interventional procedures," *J. Ultrasound Med.*, vol. 38, no. 10, pp. 2541–2557, Oct. 2019.

[19] C. P. Nolsøe et al., "Use of ultrasound contrast agents in relation to percutaneous interventional procedures: A systematic review and pictorial essay," *J. Ultrasound Med.*, vol. 37, no. 6, pp. 1305–1324, Jun. 2018.

[20] J. M. DeWitt et al., "Interventional endoscopic ultrasound: Current status and future directions," *Clin. Gastroenterol. Hepatol.*, vol. 19, no. 1, pp. 24–40, Jan. 2021.

[21] M. Antico et al., "Ultrasound guidance in minimally invasive robotic procedures," *Med. Image Anal.*, vol. 54, pp. 149–167, May 2019.

[22] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Catheter localization in 3D ultrasound using voxel-of-interest-based ConvNets for cardiac intervention," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 6, pp. 1069–1077, Jun. 2019.

[23] H. Yang, C. Shan, T. Tan, A. F. Kolen, and P. H. N. de With, "Transferring from ex-vivo to in-vivo: Instrument localization in 3D cardiac ultrasound using pyramid-UNet with hybrid loss," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2019* (Lecture Notes in Computer Science), vol. 11768. Cham, Switzerland: Springer, 2019, pp. 263–271.

[24] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Automated catheter localization in volumetric ultrasound using 3D patch-wise U-Net with focal loss," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1346–1350.

[25] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Catheter detection in 3D ultrasound using triplanar-based convolutional neural networks," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 371–375.

[26] H. Yang, C. Shan, A. Bouwman, A. F. Kolen, and P. H. N. de With, "Efficient and robust instrument segmentation in 3D ultrasound using patch-of-interest-FuseNet with hybrid loss," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101842.

[27] H. Yang, C. Shan, A. Bouwman, L. R. C. Dekker, A. F. Kolen, and P. H. N. de With, "Medical instrument segmentation in 3D U.S. by hybrid constrained semi-supervised learning," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 2, pp. 762–773, Feb. 2022.

[28] H. Yang, C. Shan, A. F. Kolen, and P. H. N. D. With, "Weakly-supervised learning for catheter segmentation in 3D frustum ultrasound," *Computerized Med. Imag. Graph.*, vol. 96, Mar. 2022, Art. no. 102037.

[29] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Deep Q-network-driven catheter segmentation in 3D U.S. by hybrid constrained semi-supervised learning and dual-UNet," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2020* (Lecture Notes in Computer Science), vol. 12261. Cham, Switzerland: Springer, 2020, pp. 646–655.

[30] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Efficient catheter segmentation in 3D cardiac ultrasound using slice-based FCN with deep supervision and F-score loss," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 260–264.

[31] H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Improving catheter segmentation localization in 3D cardiac ultrasound using direction-fused FCN," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1122–1126.

[32] L. Min, H. Yang, C. Shan, A. F. Kolen, and P. H. N. de With, "Feasibility study of catheter segmentation in 3D Frustum ultrasounds by DCNN," *Proc. SPIE*, vol. 11315, Mar. 2020, Art. no. 1131521.

[33] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, (1998), "Multiscale vessel enhancement filtering," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*. Berlin, Germany: Springer, 1998, pp. 130–137.

[34] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2016*, vol. 9901, S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells, Eds. Cham, Switzerland: Springer, 2016, pp. 424–432.

[35] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1725–1732.

[36] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[37] S. Chen, K. Ma, and Y. Zheng, "Med3D: Transfer learning for 3D medical image analysis," 2019, *arXiv:1904.00625*.

[38] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2921–2929.

[39] Y. Zhang et al., "Multi-needle localization with attention U-Net in U.S.-guided HDR prostate brachytherapy," *Med. Phys.*, vol. 47, no. 7, pp. 2735–2745, Jul. 2020.

[40] Y. Zhang et al., "Automatic multi-needle localization in ultrasound images using large margin mask RCNN for ultrasound-guided prostate brachytherapy," *Phys. Med. Biol.*, vol. 65, no. 20, Oct. 2020, Art. no. 205003.

[41] C. Andersén, T. Rydén, P. Thunberg, and J. H. Lagerlöf, "Deep learning-based digitization of prostate brachytherapy needles in ultrasound images," *Med. Phys.*, vol. 47, no. 12, pp. 6414–6420, Dec. 2020.

[42] F. Wang, L. Xing, H. Bagshaw, M. Buyyounouski, and B. Han, "Deep learning applications in automatic needle segmentation in ultrasound-guided prostate brachytherapy," *Med. Phys.*, vol. 47, no. 9, pp. 3797–3805, Sep. 2020.

[43] D. Liu et al., "The challenges facing deep learning–based catheter localization for ultrasound guided high-dose-rate prostate brachytherapy," *Med. Phys.*, vol. 49, no. 4, pp. 2442–2451, Apr. 2022.

[44] K. B. Girum, A. Lalande, R. Hussain, and G. Créhange, "A deep learning method for real-time intraoperative U.S. image segmentation in prostate brachytherapy," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 9, pp. 1467–1476, Sep. 2020.

[45] K. B. Girum, G. Créhange, R. Hussain, and A. Lalande, "Fast interactive medical image segmentation with weakly supervised deep learning method," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 9, pp. 1437–1444, Sep. 2020.

[46] N. Orlando, D. J. Gillies, I. Gyacskov, C. Romagnoli, D. D'Souza, and A. Fenster, "Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images," *Med. Phys.*, vol. 47, no. 6, pp. 2413–2426, Jun. 2020.

[47] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Proc. Int. Workshop Deep Learn. Med. Image Anal.*, vol. 11045, 2018, pp. 3–11.

[48] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Trans. Med. Imag.*, vol. 39, no. 6, pp. 1856–1867, Jun. 2020.

[49] N. Orlando et al., "Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound," *Phys. Med. Biol.*, vol. 67, no. 7, Apr. 2022, Art. no. 074002.

[50] Y. Lei et al., "Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net," *Med. Phys.*, vol. 46, no. 7, pp. 3194–3206, Jul. 2019.

[51] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.

[52] E. M. A. Anas et al., "Clinical target-volume delineation in prostate brachytherapy using residual neural networks," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2017*, vol. 10435, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 365–373.

[53] D. Karimi et al., "Accurate and robust segmentation of the clinical target volume for prostate brachytherapy," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2018*, vol. 11073, A. Frangi, J. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, Eds. Cham, Switzerland: Springer, 2018, pp. 531–539.

[54] D. Karimi et al., "Accurate and robust deep learning-based segmentation of the prostate clinical target volume in ultrasound images," *Med. Image Anal.*, vol. 57, pp. 186–196, Oct. 2019.

[55] X. He et al., "Deep attentional GAN-based high-resolution ultrasound imaging," *Proc. SPIE*, vol. 11319, Mar. 2020, Art. no. 113190B.

[56] M. Golshan et al., "Automatic detection of brachytherapy seeds in 3D ultrasound images using a convolutional neural network," *Phys. Med. Biol.*, vol. 65, no. 3, Feb. 2020, Art. no. 035016.

[57] Y. LeCun et al., "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[58] Y. Chen et al., "MR to ultrasound image registration with segmentation-based learning for HDR prostate brachytherapy," *Med. Phys.*, vol. 48, no. 6, pp. 3074–3083, Jun. 2021.

[59] Q. Zeng et al., "Weakly non-rigid MR-TRUS prostate registration using fully convolutional and recurrent neural networks," *Proc. SPIE*, vol. 11313, Mar. 2020, Art. no. 113132Y.

[60] J. R. Rodgers, D. J. Gillies, W. T. Hrinivich, I. Gyackov, and A. Fenster, "Automatic needle localization in intraoperative 3D transvaginal ultrasound images for high-dose-rate interstitial gynecologic brachytherapy," *Proc. SPIE*, vol. 11315, Mar. 2020, Art. no. 113150K.

[61] H. Sun, K. Zhang, R. Fan, W. Xiong, and J. Yang, "Stepwise local synthetic pseudo-CT imaging based on anatomical semantic guidance," *IEEE Access*, vol. 7, pp. 168428–168435, 2019.

[62] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.

[63] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (Still) use DBSCAN," *ACM Trans. Database Syst.*, vol. 42, no. 3, pp. 1–21, Sep. 2017.

[64] N. Abraham and N. M. Khan, "A novel focal Tversky loss function with improved attention U-Net for lesion segmentation," 2018, *arXiv:1810.07842*.

[65] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[66] W. T. Hrinivich et al., "Simultaneous automatic segmentation of multiple needles using 3D ultrasound for high-dose-rate prostate brachytherapy," *Med. Phys.*, vol. 44, no. 4, pp. 1234–1245, Apr. 2017.

[67] D. Viderman, M. Dossov, S. Seitenov, and M.-H. Lee, "Artificial intelligence in ultrasound-guided regional anesthesia: A scoping review," *Frontiers Med.*, vol. 9, p. 3037, Oct. 2022.

[68] C. Mwikirize, J. L. Nosher, and I. Hacihaliloglu, "Convolution neural networks for real-time needle detection and localization in 2D ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 5, pp. 647–657, May 2018.

[69] C. Mwikirize, J. L. Nosher, and I. Hacihaliloglu, "Single shot needle tip localization in 2D ultrasound," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2019*. (Lecture Notes in Computer Science), vol. 11768. Cham, Switzerland: Springer, 2019, pp. 637–645.

[70] C. Mwikirize, A. B. Kimbowa, S. Imanirakiza, A. Katumba, J. L. Nosher, and I. Hacihaliloglu, "Time-aware deep neural networks for needle tip localization in 2D ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 5, pp. 819–827, May 2021.

[71] J. Gao, P. Liu, G.-D. Liu, and L. Zhang, "Robust needle localization and enhancement algorithm for ultrasound by deep learning and beam steering methods," *J. Comput. Sci. Technol.*, vol. 36, no. 2, pp. 334–346, Apr. 2021.

[72] A. Pourtaherian et al., "Improving needle detection in 3D ultrasound using orthogonal-plane convolutional networks," *Medical Image Computing and Computer-Assisted Intervention MICCAI 2017*, vol. 10434, M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins, and S. Duchesne, Eds. Cham, Switzerland: Springer, 2017, pp. 610–618.

[73] A. Pourtaherian et al., "Localization of partially visible needles in 3D ultrasound using dilated CNNs," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2018, pp. 1–4.

[74] E. Smistad and F. Lindseth, "Real-time automatic artery segmentation, reconstruction and registration for ultrasound-guided regional anaesthesia of the femoral nerve," *IEEE Trans. Med. Imag.*, vol. 35, no. 3, pp. 752–761, Mar. 2016.

[75] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. G. van Sloun, and Y. C. Eldar, "KalmanNet: Neural network aided Kalman filtering for partially known dynamics," *IEEE Trans. Signal Process.*, vol. 70, pp. 1532–1547, 2022.

[76] A. Arjas et al., "Neural network Kalman filtering for 3-D object tracking from linear array ultrasound data," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 5, pp. 1691–1702, May 2022.

[77] E. Smistad, K. F. Johansen, D. H. Iversen, and I. Reinertsen, "Highlighting nerves and blood vessels for ultrasound-guided axillary nerve block procedures using neural networks," *J. Med. Imag.*, vol. 5, no. 4, Nov. 2018, Art. no. 044004.

[78] M. Baby and A. S. Jereesh, "Automatic nerve segmentation of ultrasound images," in *Proc. Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Apr. 2017, pp. 107–112.

[79] H. M. Díaz-Vargas, C. A. Jimenez-Castaño, D. A. Cárdenas-Peña, O. D. Aguirre-Ospina, and A. A. Orozco-Gutierrez, "Peripheral nerve segmentation in ultrasound images using conditioned U-Net," in *Progress in Artificial Intelligence and Pattern Recognition* (Lecture Notes in Computer Science), vol. 13055, Y. H. Heredia, V. M. Núñez, and J. R. Shulcloper, Eds. Cham, Switzerland: Springer, 2021, pp. 124–133.

[80] E. Maneas et al., "Deep learning for instrumented ultrasonic tracking: From synthetic training data to in vivo application," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 69, no. 2, pp. 543–552, Feb. 2022.

[81] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[82] C. Mwikirize, J. L. Nosher, and I. Hacihaliloglu, "Learning needle tip localization from digital subtraction in 2D ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 6, pp. 1017–1026, Jun. 2019.

[83] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*, vol. 9351, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[84] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Apr. 2016, pp. 770–778.

[85] H. Rivaz et al., "Ablation monitoring with elastography: 2D in-vivo and 3D ex-vivo studies," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2008*, vol. 5242, D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds. Berlin Cham, Switzerland: Springer, 2008, pp. 458—466.

[86] S. Zhang et al., "Detection and monitoring of thermal lesions induced by microwave ablation using ultrasound imaging and convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 965–973, Apr. 2020.

[87] M. Wang et al., "Detection and monitoring of microwave ablation by ultrasound imaging based on convolutional neural network," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2020, pp. 1–3.

[88] R. Kondo, N. Koizumi, Y. Nishiyama, N. Matsumoto, and K. Numata, "Out-of-plane motion detection system using convolutional neural network for U.S.-guided radiofrequency ablation therapy," in *Proc. 15th Int. Conf. Ubiquitous Robots (UR)*, Jun. 2018, pp. 729–731.

[89] M. Arif, A. Moelker, and T. van Walsum, "Automatic needle detection and real-time bi-planar needle visualization during 3D ultrasound scanning of the liver," *Med. Image Anal.*, vol. 53, pp. 104–110, Apr. 2019.

[90] L. Mercier et al., "Online database of clinical MR and ultrasound images of brain tumors," *Med. Phys.*, vol. 39, p. 3253, Jun. 2012.

[91] Y. Xiao et al., "Evaluation of MRI to ultrasound registration methods for brain shift correction: The CuRIOUS2018 challenge," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 777–786, Mar. 2020.

[92] Y. Xiao, M. Fortin, G. Unsgard, H. Rivaz, and I. Reinertsen, "REtro-Spective evaluation of cerebral tumors (RESECT): A clinical database of pre-operative MRI and intra-operative ultrasound in low-grade glioma surgeries," *Med. Phys.*, vol. 44, no. 7, pp. 3875–3882, Jul. 2017, doi: 10.1002/mp.12268.

[93] L. Canalini, J. Klein, D. Miller, and R. Kikinis, "Segmentation-based registration of ultrasound volumes for glioma resection in image-guided neurosurgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 10, pp. 1697–1713, Oct. 2019.

[94] L. Canalini, J. Klein, D. Miller, and R. Kikinis, "Enhanced registration of ultrasound volumes by segmentation of resection cavity in neurosurgical procedures," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 12, pp. 1963–1974, Dec. 2020.

[95] R. A. Zeineldin, M. E. Karar, J. Coburger, C. R. Wirtz, F. Mathis-Ullrich, and O. Burgert, "Towards automated correction of brain shift using deep deformable magnetic resonance imaging-intraoperative ultrasound (MRI-iUS) registration," *Current Directions Biomed. Eng.*, vol. 6, no. 1, pp. 1–5, Sep. 2020.

[96] R. A. Zeineldin et al., "IRegNet: Non-rigid registration of MRI to interventional U.S. for brain-shift compensation using convolutional neural networks," *IEEE Access*, vol. 9, pp. 147579–147590, 2021.

[97] R. A. Zeineldin, M. E. Karar, F. Mathis-Ullrich, and O. Burgert, "A hybrid deep registration of MR scans to interventional ultrasound for neurosurgical guidance," in *Machine Learning in Medical Imaging*, vol. 12966, C. Lian, X. Cao, I. Rekik, X. Xu, and P. Yan, Eds. Cham, Switzerland: Springer, 2021, pp. 586–595.

[98] A. Pirhadi, S. Salari, M. O. Ahmad, H. Rivaz, and Y. Xiao, "Robust landmark-based brain shift correction with a Siamese neural network in ultrasound-guided brain tumor resection," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 18, pp. 501–508, Oct. 2022.

[99] M. Cen and C. Jung, "Fully convolutional Siamese fusion networks for object tracking," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 850–865.

[100] R. A. Zeineldin, A. Pollok, T. Mangliers, M. E. Karar, F. Mathis-Ullrich, and O. Burgert, "Deep automatic segmentation of brain tumours in interventional ultrasound data," *Current Directions Biomed. Eng.*, vol. 8, no. 1, pp. 133–137, Jul. 2022.

[101] J. Chen et al., "TransUNet: Transformers make strong encoders for medical image segmentation," 2021, *arXiv:2102.04306*.

[102] F. X. Carton, M. Chabanas, B. K. Munkvold, I. Reinertsen, and J. H. Noble, "Automatic segmentation of brain tumor in intraoperative ultrasound images using 3D U-Net," *Proc. SPIE*, vol. 11315, pp. 190–195, Mar. 2020.

[103] S. Cepeda et al., 'Comparison of intraoperative ultrasound B-mode and strain elastography for the differentiation of glioblastomas from solitary brain metastases. An automated deep learning approach for image analysis," *Frontiers Oncol.*, vol. 10, Feb. 2021, Art. no. 590756.

[104] M. P. Heinrich and L. Hansen, "Highly accurate and memory efficient unsupervised learning-based discrete CT registration using 2.5D displacement search," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 190–200.

[105] J. H. Lee et al., "Deep learning with ultrasonography: Automated classification of liver fibrosis using a deep convolutional neural network," *Eur. Radiol.*, vol. 30, no. 2, pp. 1264–1273, Feb. 2020.

[106] D. J. Gillies et al., "Deep learning segmentation of general interventional tools in two-dimensional ultrasound images," *Med. Phys.*, vol. 47, no. 10, pp. 4956–4970, Oct. 2020.

[107] Y. Wang et al., "Deep attentive features for prostate segmentation in 3D transrectal ultrasound," *IEEE Trans. Med. Imag.*, vol. 38, no. 12, pp. 2768–2778, Dec. 2019.

[108] Y. Hu et al., "Adversarial deformation regularization for training image registration neural networks," in *Medical Image Computing and Computer Assisted Intervention MICCAI 2018*, vol. 11070. Cham, Switzerland: Springer, 2018, pp. 774–782.

[109] T. Leblanc, F. Lalys, Q. Tollenaere, A. Kaladji, A. Lucas, and A. Simon, "Stretched reconstruction based on 2D freehand ultrasound for peripheral artery imaging," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, no. 7, pp. 1281–1288, Jul. 2022.

[110] M. Luo, "Deep motion network for freehand 3D ultrasound reconstruction," in *Proc. MICCAI*, vol. 13434. Cham, Switzerland: Springer, 2022, pp. 290–299.

[111] H. Guo, S. Xu, B. Wood, and P. Yan, "Sensorless freehand 3D ultrasound reconstruction via deep contextual learning," in *Proc. MICCAI*, 2020, pp. 463–472.

[112] P. U. Pandey, N. Quader, P. Guy, R. Garbi, and A. J. Hodgson, "Ultrasound bone segmentation: A scoping review of techniques and validation practices," *Ultrasound Med. Biol.*, vol. 46, no. 4, pp. 921–935, Apr. 2020.

[113] I. Hacihaliloglu, "Ultrasound imaging and segmentation of bone surfaces: A review," *Technology*, vol. 5, no. 2, pp. 74–80, Jun. 2017.

[114] S. Schumann, "State of the art of ultrasound-based registration in computer assisted orthopedic interventions," in *Computational Radiology for Orthopaedic Interventions*, vol. 23, G. Zheng and S. Li, Eds. Cham, Switzerland: Springer, 2016, pp. 271–297.

[115] G. Ayana, K. Dese, and S.-W. Choe, "Transfer learning in breast cancer diagnoses via ultrasound imaging," *Cancers*, vol. 13, no. 4, p. 738, Feb. 2021.

[116] A. Hering et al., "Learn2Reg: Comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning," *IEEE Trans. Med. Imag.*, vol. 42, no. 3, pp. 697–712, Mar. 2023.

[117] M. Antonelli, A. Reinke, and S. Bakas, "The medical segmentation decathlon," *Nature Commun.*, vol. 13, p. 4128, Jul. 2022.

[118] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM Comput. Surv.*, vol. 54, no. 10s, pp. 1–41, Jan. 2022.

[119] T. Elsken, J. H. Metzen, and F. Hutter, "Neural architecture search: A survey," *J. Mach. Learn. Res.*, vol. 20, no. 1, pp. 1997–2017, 2019.

[120] A. Rasoulian, S. Salari, and Y. Xiao, "Weakly supervised intracranial hemorrhage segmentation using hierarchical combination of attention maps from a Swin transformer," in *Machine Learning in Clinical Neuroimaging* (Lecture Notes in Computer Science), vol. 13596. Cham, Switzerland: Springer, 2022, pp. 63–72.

[121] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 618–626.

[122] D. Minh, H. X. Wang, Y. F. Li, and T. N. Nguyen, "Explainable artificial intelligence: A comprehensive review," *Artif. Intell. Rev.*, vol. 55, no. 5, pp. 3503–3568, Jun. 2022.