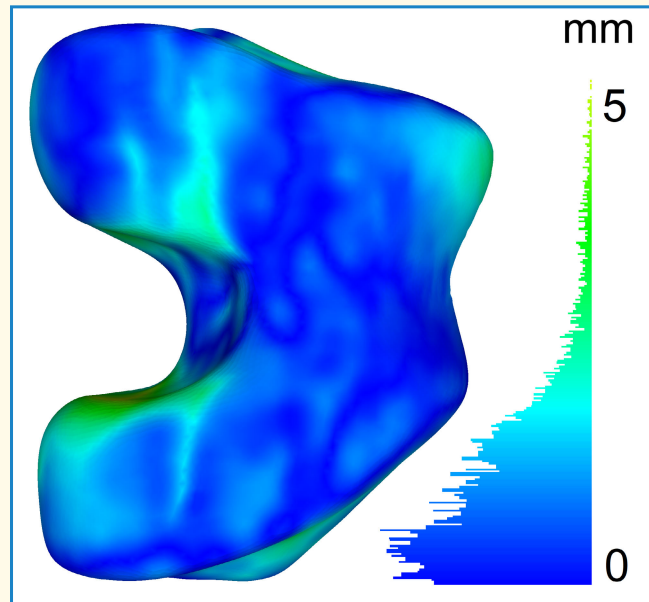


# Knee Bone Models From Ultrasound

Benjamin Hohlmann<sup>1</sup>, Peter Broessner<sup>1</sup>, Lovis Phlippen, Thorsten Rohde, and Klaus Radermacher

**Abstract**—The number of total knee arthroplasties performed worldwide is on the rise. Patient-specific planning and implants may improve surgical outcomes but require 3-D models of the bones involved. Ultrasound (US) may become a cheap and nonharmful imaging modality if the shortcomings of segmentation techniques in terms of automation, accuracy, and robustness are overcome; furthermore, any kind of US-based bone reconstruction must involve some kind of model completion to handle occluded areas, for example, the frontal femur. A fully automatic and robust processing pipeline is proposed, generating full bone models from 3-D freehand US scanning. A convolutional neural network (CNN) is combined with a statistical shape model (SSM) to segment and extrapolate the bone surface. We evaluate the method in vivo on ten subjects, comparing the US-based model to a magnetic resonance imaging (MRI) reference. The partial freehand 3-D record of the femur and tibia bones deviate by 0.7–0.8 mm from the MRI reference. After completion, the full bone model shows an average submillimetric error in the case of the femur and 1.24 mm in the case of the tibia. Processing of the images is performed in real time, and the final model fitting step is computed in less than a minute. It took an average of 22 min for a full record per subject.

**Index Terms**—Convolutional neural network (CNN), knee, model completion, segmentation, statistical shape model (SSM), ultrasound (US).



## I. INTRODUCTION

REPLACING the knee joint surface with an implant is one of the most frequently conducted surgeries worldwide [1]. Statistically, one in four people in the United States will receive a total knee arthroplasty (TKA) in their lifetime [2], [3], [4]. The main indication for knee replacement is osteoarthritis, which affects 13% of women and 10% of men over the age of 60 in the United States [5].

The shape, size, and positioning of the implant are crucial factors for a functional knee replacement and acceptable patient satisfaction [6], [7]. Different anatomical structures require different implant sizes, whereas a linear scaling of a one-shape implant is not sufficient for a satisfactory result [8],

Manuscript received 27 April 2023; accepted 6 June 2023. Date of publication 22 June 2023; date of current version 29 August 2023. (Corresponding author: Benjamin Hohlmann.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethikkommission an der Medizinischen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen under Application No. 22-397.

Benjamin Hohlmann, Peter Broessner, Lovis Phlippen, and Klaus Radermacher are with the Chair for Medical Engineering, RWTH Aachen University, 52062 Aachen, Germany (e-mail: hohlmann@hia.rwth-aachen.de).

Thorsten Rohde is with the Clinic for Trauma Surgery and Orthopedics, Luisenhospital Aachen, 52064 Aachen, Germany.

Digital Object Identifier 10.1109/TUFFC.2023.3286287

[9], [10]. Individual implant shapes have been developed with the motivation of achieving optimal results for each patient [11]. Studies show advantageous factors of patient-specific implants compared to off-the-shelf implants, such as reduced blood loss during surgery, fewer adverse event rates, and lower rates of being discharged to rehabilitation facilities. In comparison to off-the-shelf implants, the total costs of care for patient-specific implants are even lower [12], [13]. Another impact on the outcome is the implant placement, and different approaches exist: navigation [14], robotics [15], and patient-specific instrumentation [16], [17], [18]. A previous image acquisition of the knee is required for patient-specific instrumentation and implants in order to generate 3-D bone models for surgical planning.

According to the literature, 3-D planning has advantages even for off-the-shelf implants, for example, for correct rotation of the femoral components [19], [20]. The state of the art for acquiring 3-D bone models is computed tomography (CT). The CT scans of current systems achieve an in-plane voxel spacing of 0.3–0.5 mm [21]. The drawbacks of CT are exposure of the patient to radiation and high costs. The CT scans may be associated with an increased risk of cancer [22], and the costs are approximately ten times higher for a CT scan than for an ultrasound (US) examination, according to health insurance billing lists in Germany [23].

### Highlights

- **The first in-vivo evaluation in a near clinical setting of a real-time capable method for generating knee bone models from ultrasound.**
- **The proposed methods can reconstruct the surface of the distal femur and proximal tibia with an average surface distance error of 0.96 and 1.24 mm, respectively.**
- **Ultrasound may become a cheap and widely available point-of-care imaging alternative to computed tomography and magnetic resonance imaging in orthopedics.**

Apart from lower costs, US as an alternative medical imaging technique has the advantage of being nonharmful [24]. Moreover, it offers the possibility of visualizing various tissues or bones and real-time capability, which enables functional diagnosis [21], [25]; however, US also has some disadvantages: noise due to diffuse deflections, imaging artifacts, blurred bone surfaces, and bone shadowing [26], [27]. Probes for acquiring 2-D images can be combined with a localizer system, which can be optical, electromagnetic, or mechanical, for the 3-D imaging of large anatomical areas, such as the knee [26]. This approach is called “freehand 3-D US.” Given the drawbacks mentioned above, every knee bone US scan will lack information on the inner tibiofemoral and patellofemoral joint space, which is occluded by the patella. As such, some kind of model completion is necessary to provide full bone models to the clinician.

## II. RELATED WORK

The reconstruction of complete bone surface models from US data is mostly approached by a two-stage concept: first, US images are segmented for the derivation of a partial surface point cloud. Second, the partial surface scan is completed to a full bone model using statistical shape models (SSMs). Accordingly, we first present related work on US image segmentation, followed by related work on US-based model completion.

### A. US Image Segmentation

Bone surfaces need to be segmented in US images for the extraction of partial bone surface points. In addition to characteristics of US images, such as noise and artifacts, one of the most challenging aspects of US segmentation is the distinction of bone interfaces from other tissue interfaces. The survey of Pandey et al. [25] provides a comprehensible introduction to the topic and gives an overview of US segmentation to the year 2019.

Ronneberger et al. [28] established a standard for medical image segmentation in 2015 based on learned features. They introduce a convolutional neural network (CNN)-based encoder-decoder architecture with skip connections called “U-Net.” Subsequent publications introduce extensions, such as the 3-D U-Net [29], the U-Net++ with redesigned skip connections [30] or the nnU-Net framework for the automated configuration of architecture and training processes [31].

Most works on bone segmentation in US images rely on the original U-Net [32], [33], [34], [35], [36], [37], [38], [39], [40], [41] and the U-Net++ [42]. Several works introduce

dropout during inference [43], [44] or training [45] to incorporate segmentation uncertainty. Other approaches use auxiliary hand-crafted features, such as phase symmetry or bone shadow [46], [47], [48]. Banerjee et al. [49] extended the U-Net with depthwise separable convolution and residual connections for their LDS-U-Net, similar to [50], who added depthwise separable convolution next to an attention mechanism in their BoneNet. Rahman et al. [51] introduce graph convolution to a U-Net-like architecture in order to use the characteristic connectivity of bone surfaces in US images. Dunnhofer et al. [52] used the information of neighboring slices using two encoders in a Siam-U-Net architecture for the segmentation of volumetric US images.

In addition to U-Net and its variants, other architectures employed for US bone segmentation include the Mask R-CNN [53], Pyramid Attention Network [54], and DeepLabv3+ [55], [56], [57]; further approaches include the use of conditional generative adversarial networks, such as the pix2pix architecture introduced by Isola et al. [58] for image-to-image translation. Zhou et al. [59] employ pix2pix for direct segmentation and the improvement of U-Net segmentations. Similarly, Alsinan et al. [60] condition a conditional generative adversarial network for the generation of bone shadow maps, which are then fused with CNN-based features.

### B. Reconstruction

The SSMs allow the incorporation of morphological statistics for the creation of new shapes as a linear combination of mean shape  $\bar{x}$  and modes matrix  $P$ . Approaches for US-based model completion to date rely mostly on SSMs for shape generation but differ in the methods of building the SSM and adapting it to US data. Morooka et al. [61] and Sarkalkan et al. [62] give an overview of US-based model completion; we refer to [63] for a survey and comprehensible introduction.

Hacihaliloglu et al. [64] propose a statistical model of the shape and pose of the spine and fit it to local phase features extracted from 3-D US data for reconstruction. Regarding model adaption, they employ a method based on Gaussian mixture models combined with an expectation-maximization scheme. Klicken oder tippen Sie hier, um Text einzugeben. or a quasi-Newton optimization [65], [66], which bypasses the need for correspondences.

Anas et al. [67] propose a very similar concept, using a statistical model of pose and shape for the intraoperative reconstruction of carpal bones from US. They, first adapt pose and shape components to preoperative CT data, and then only

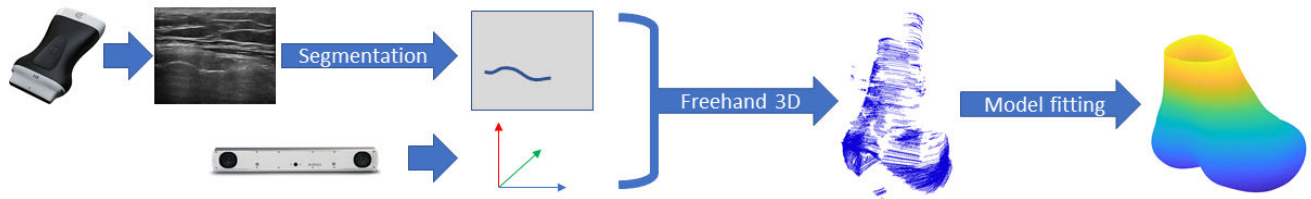


Fig. 1. General workflow of the proposed method. US images are acquired with a wireless probe, segmented by a CNN, and combined with tracking data to a partial point cloud of the bone. The SSM is then fit to the partial scan to obtain a full model of the target bone.

adapt pose components for the subsequent reconstruction from segmented US images. The implementation of the model adaptation is again based on correspondence-free Gaussian mixture models, initialized with landmark-based registration.

As opposed to previous methods, some approaches rely on point correspondences for SSM adaption to US: Hänisch et al. [68] adapt an SSM to segmented 3-D US images for the reconstruction of knee bones. They employ a variant of the iterative closest point algorithm for establishing point correspondences. In a subsequent publication, correspondences are established by searching for high-intensity pixels along vertex normals in the original 3-D US image [54], [69].

The recent work of Mahfouz et al. [70] is most closely related to our work: they employ 3-D freehand US with electromagnetic tracking for image acquisition. Bone surfaces are segmented in the frequency-range data acquired and combined into a 3-D partial point cloud based on tracking information. After the outlier removal on the partial point cloud, they match it to an SSM for the completion of the distal femoral and proximal tibial bone models; however, essential details, such as the fitting method of their implementation, remain unknown.

Compared to related work, the objectives of our work are as follows.

- 1) The development of a new SSM adoption algorithm based on a general-purpose optimizer (GPO).
- 2) The evaluation of the GPO approach against the state of the art.
- 3) The proposal of a fully automated workflow with real-time capability.
- 4) The presentation of an entire pipeline, including all the algorithms.
- 5) Open access to test samples and results.
- 6) The first in vivo evaluation in a near-clinical setting.

### III. METHODS

The workflow proposed is a multistep procedure, as demonstrated in Fig. 1: landmarks were acquired using a freehand 3-D setup with a 2-D probe and a tracking system for the prepositioning of the image volume relative to the mean shape of the SSM. Next, two partial scans of the knee were acquired using the US probe. The 2-D images were segmented fully automatically in real time by a CNN. After the acquisition, the partial scans of the knee were combined with the SSM to reconstruct a full bone model of the proximal femur and distal tibia.

#### A. Experimental Setup

We used a Clarius L15 HD (Clarius, Vancouver, BC, Canada) probe to acquire the US images. The Clarius system



Fig. 2. Experimental setup. The subject's leg is fixated by an adjustable boot. The subject's leg and the wireless probe are tracked by an optical tracking system. Our custom user interface shows the current US image and the recorded 3-D point cloud. Left: anterior scan with the subject in a sitting position and 90° knee flexion. Right: posterior scan with the subject in a prone position and the knee in extension.

provides an advanced programming interface, which we use to feed the image data into our custom software. The imaging depth was adjusted according to the approximate depth of the bone. The Clarius software provided fully automatic adjustment of imaging parameters, such as the frequency, easing the acquisition process.

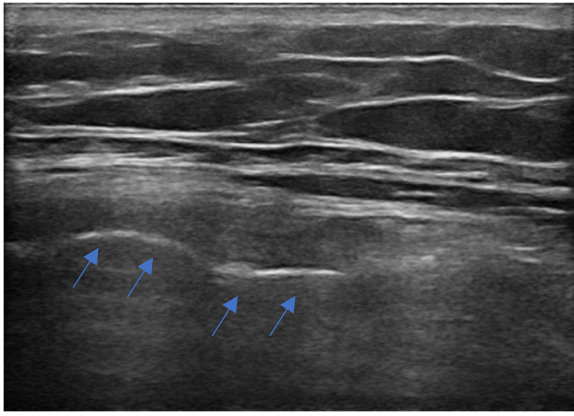
The FusionTrack500 (Atracsys, Puidoux, Switzerland) localizer system, which comes with its own advanced programming interface, was used for tracking. We tracked the position of the probe and a dynamic reference base attached to the subject's lower leg by Velcro braces at 300 Hz. Tracking data were matched with image data based on the acquisition time recorded by our software. The beamforming and image transfer process of the US images induced a latency, which we determined in preliminary experiments to be 175 ms. Given the time stamps, we matched the image and tracking data and constructed the following transformation chain:

$$T_{\text{Total}} = T_{\text{model}} * T_{\text{DRB}}^{-1} * T_{\text{Probe}} * T_{\text{Calib}}$$

where  $T_{\text{Calib}}$  was determined from computer-aided design data of the 3-D-printed adapter connecting the tracking body to the probe and validated using a sphere-based method.  $T_{\text{Probe}}$  was the pose of the probe,  $T_{\text{DRB}}$  was the pose of the dynamic reference base and  $T_{\text{model}}$  was the transformation of the anatomical landmarks on the mean shape of the SSM to the anatomical landmarks of the subject acquired prior to the record.

The subject was put in a sitting position with the leg in 90° flexion. The patella is moved distally in this pose, minimizing the occlusion of the femur. The optical tracking camera was positioned at the subject's feet area. The clinical expert sonographer was standing to the side of the subject, enabling free movement and providing a view of the knee and our custom software simultaneously (see Fig. 2). The approximate





**Fig. 3.** One of the US images used for training the CNN architecture. Note how soft-tissue and bone interfaces resemble each other. The actual bone surface is just the lowest bright ridge indicated by the blue arrows.

position of the femur was determined by acquiring the anatomical landmarks, namely, the medial and lateral epicondyle, as well as a point on the distal femoral shaft. The tibial landmarks were medial and lateral points on the proximal crest, as well as a point on the proximal shaft. Note that the rather ambiguous points on the shafts were chosen due to the requirement that all points are accessible in both the prone and seated position. Alternative registration methods include pair-wise surface-based registration and registration onto the SSM mean shape. After the frontal femur and tibia had been recorded, the subject was asked to lie down in a prone position, and the scanning of the posterior femur and tibia was performed. During this acquisition process, the sonographer tried to optimize the probe orientation to the bone surface in order to obtain a distinct bone response and tried not to include bones other than the respective target anatomy in the record. After the record had been completed, both point sets were registered to the SSM mean shape using a custom MATLAB (MathWorks, Natick, MA, USA) implementation of the iterative closest point algorithm that is robust to outliers. The point sets acquired by this process were still partial point sets, as the inner parts of the knee joint spaces cannot be imaged due to bone shadowing.

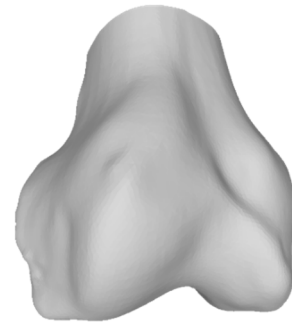
## B. Datasets

Prior to the experiment described above, training data for the individual pipeline components were acquired as follows.

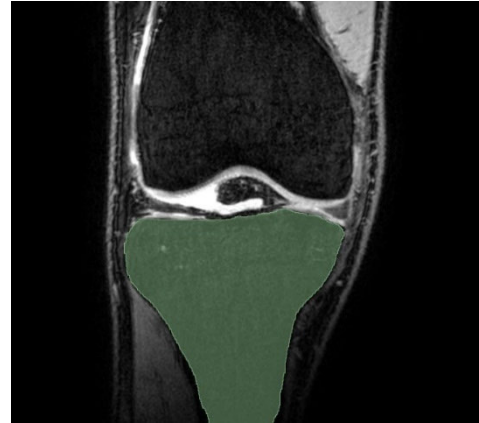
**1) US Segmentation:** A separate set of 4565 images was acquired and annotated beforehand. The images depict all parts of the knee joint, including areas showing no bone surface at all. Only the thin bone response was annotated and not the bone shadow. The dataset included ten subjects, ages 26–38, seven male and three female.

The dataset was split 10:1 into a train and a validation dataset. Image height varied from 150 to 750 pixels, and image width 420–650 pixels. Accordingly, the pixel spacing varied from 0.08 to 0.12 mm per pixel. See Fig. 3 for an example image.

**2) Statistical Shape Model:** Two SSMs of the tibia and femur, respectively, were built from 414 anonymized bone



**Fig. 4.** Segmented femur of one of the 414 surface models used to build the SSM.



**Fig. 5.** One slice of the MRI of one of the subjects. The manual ground truth annotation of the tibia is overlaid.

surface datasets of patients that had undergone TKA. All patients underwent CT imaging, which was subsequently manually segmented and freed from osteophytes by experts. The mean shape surface models were aligned according to anatomical landmarks. See Fig. 4 for an example mesh of the distal femur. We refer the reader to our previous publication for additional details [54].

**3) Reference: MRI:** Magnetic resonance imaging (MRI) was used as a reference to evaluate the accuracy of the US-based reconstruction. Nine subjects of age 28–58 were recruited, five males and four females. A 3-D water-selective cartilage scan was recorded, and the femur and tibia were manually segmented by medical engineering students. The segmentation was validated by an expert with more than five years of clinical experience. All volumes have an image size of  $720 \times 720 \times 266$ . The slice thickness is 0.75 mm, and the in-plane pixel spacing is 0.25 mm. See Fig. 5 for an example image and its annotation performed with 3-DSlicer [71]. A recent high-resolution CT image was available for one subject.

## C. Segmentation by CNN

We opted for the Deeplabv3+ for image segmentation as it combines a fast processing speed with high accuracy. The lightweight MobileNetV2 backbone was chosen as its low model capacity enabled better generalization of our small dataset. At the same time, this enabled fast inference. We refer

to our previous publication [57] and the article by Chen et al. [72] for implementation details. We did not find any benefit from employing more recent CNN or vision transformer architectures on the task of bone segmentation [57], [73]. The network was trained on a separate dataset for the experiment at hand, described in Section III-B. Hyperparameter tuning involved the learning rate and loss function. The bone surface areas inferred were skeletonized to a single pixel-wide centerline during the US image acquisition.

#### D. Bone Model Completion by SSM

All methods were implemented with MATLAB. Several algorithms were applied to adopt the SSM to the incomplete point set: 1) the active shape model (ASM) search, originally introduced by Cootes et al. [74], alternates between establishing correspondences by nearest-neighbor search and updating the rigid transformation and modes by projecting the vertex position deltas into mode space; 2) a simultaneous least-squares (LSO) optimization of pose and shape, as suggested by Blanz et al. [75]; and 3) direct optimization of pose and shape by minimizing an objective function on the average surface distance between the shape model and the incomplete point cloud. A GPO, namely, “fminsearch” by MATLAB, was used to minimize the objective function. Hyperparameter optimization, for example, regarding the number of modes predicted or the number of iterations that were tuned in preliminary *in silico* studies.

#### E. Evaluation

The performance of the methods was evaluated in terms of speed and accuracy. Computation times for segmenting 2-D image slices and the adoption of the shape models are reported, as well as total times for image acquisition of the individual bones and subjects. Regarding accuracy, the average surface distance error (SDE) directed from 1) the partially recorded point set and 2) the full reconstructed model to the reference MRI mesh is reported. See (1) for the definition, where  $X$  and  $Y$  are the source and target point sets, respectively,  $N$  the size of the source point set and  $\text{dist}(x, y)$  the Euclidean distance between two points in  $\mathbb{R}^3$

$$\text{SDE}(X, Y) = \frac{1}{N} \sum_{i=1}^N \min_{u \in Y} \text{dist}(x_i, u). \quad (1)$$

As the partial set only represents part of the bone, a directed distance from the partial scan to the ground truth mesh was computed. It should be noted that errors in terms of bone surfaces not segmented by the algorithm are not quantified by this evaluation. A directed distance from the adapted model to the ground truth mesh was computed for the full reconstruction. Only those parts of the bones that are relevant for TKA planning were evaluated (see Fig. 6). In contrast to the partial scan, the adapted model is guaranteed to include the entire bone surface of the knee.

Additionally, we provide the overall anterior-posterior (AP) and medial-lateral (ML) sizes of both the MRI reference and the US-based reconstruction of the femur and tibia.

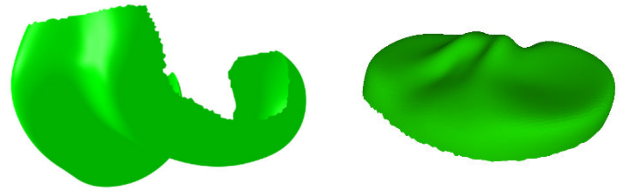


Fig. 6. Parts of the reconstructed bones on which the metrics are computed.

TABLE I

DIRECTED MEAN SURFACE DISTANCE FROM THE PARTIAL RECONSTRUCTION TO THE GROUND TRUTH, AVERAGED OVER ALL TEN SUBJECTS. EITHER JUST THE FRONT OR THE COMBINATION OF THE FRONTAL AND BACK RECORD IS REPORTED FOR EACH BONE. THE RATHER HIGH VALUES INDICATE NEIGHBORING BONES BEING VISIBLE IN THE RECORDS

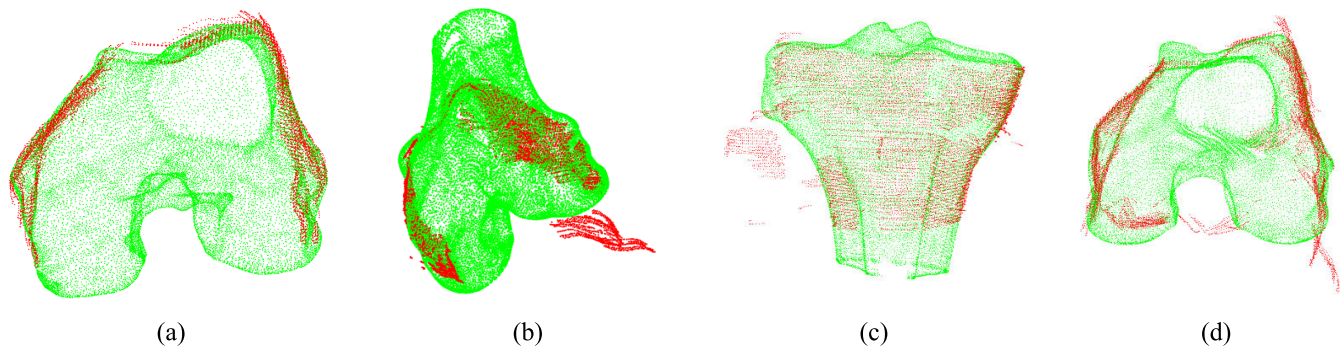
SDE (mm)	Femur	Tibia
Front only	1.19	2.27
Front & back	2.26	3.66

## IV. RESULTS

Table I presents the evaluation of the partial US records. The mean distance from the recorded frontal femoral surface to the ground truth ranges from 0.81 to 2.68 mm. The high value of 2.68 mm is due to the tibia being visible in the record, with all other values being well below. See Fig. 7(a) and (b) for two examples. In the case of the frontal tibia, the distance ranges from 1.4 to 3.83 mm, which appears to be much worse; however, the fibula is part of the scan in all records, as shown exemplarily in Fig. 7(c). Note that this does not affect the reconstruction quality, as these points are outliers and far from the reconstructed geometry. When adding the posterior scan to the partial point clouds, the error increases strongly. In the case of the femur, it ranges from 1.14 to 3.86 mm, and in the case of the tibia, from 2.49 to 6.72 mm. Again, this high error is mostly due to other bones being part of the records. Several cases, however, show a clear geometry mismatch; see, for example, the medial posterior femoral condyles in Fig. 7(d).

In order to isolate the surface mismatch from the false positive detection of neighboring bones, we manually cleaned the scan of subject #2 from the latter. Note that this step is not part of the regular pipeline. See Table II for the evaluation. Both the femur and tibia show errors <1 mm; however, the error increases when including the posterior part.

The directed mean surface distance of the masked reconstruction is reported in Tables III and IV. Because of the bone model being evaluated, the false positive errors of neighboring bones are no longer present in this evaluation. All three algorithms are evaluated on frontal and combined partial scans of the femur and tibia, respectively. In the case of the femur, the GPO method yields the best surface errors of 0.62–1.51 mm. The mean surface errors of all ten subjects is 0.96 mm. Both the LSO and the ASM methods fail to reconstruct subject #1. Their surface reconstruction errors are an average of 1.57 and 1.72 mm, respectively. A similar result is found for the tibia. Again, the GPO method yields



**Fig. 7.** US record (red) compared to the ground truth (green). (a) Average performing example record of the frontal femur. (b) Typical case where neighboring bones, in this case, the tibia, are part of the record. (c) Parts of the fibula are present. This occurs for all records of the tibia. (d) Combined front and back record of one femur. Notice the neighboring bones (bottom right) and a mismatch in the medial posterior condyle (bottom left).

**TABLE II**

SAME DISTANCE COMPUTED FOR JUST A SINGLE SUBJECT, AFTER THE MANUAL REMOVAL OF ALL NEIGHBORING BONE DETECTIONS. REGARDING ONLY THE FRONTAL SCAN, THE ERROR IS WELL BELOW 1 mm, BUT THE QUALITY DECREASES WHEN THE POSTERIOR PART IS INCLUDED. NOTE THAT THE MANUAL OUTLIER REMOVAL IS NOT PART OF THE REGULAR PIPELINE, WHERE THIS PROBLEM IS SOLVED BY THE SSM

SDE (mm)	Femur	Tibia
Front only	0.83	0.7
Front & back	1.06	1.15

**TABLE III**

SDE ERROR OF THE MASKED RECONSTRUCTION COMPARED TO THE GROUND TRUTH FOR EACH INDIVIDUAL SUBJECT USING GPO. NOTE THAT SUBJECT #4 WAS EVALUATED USING A CT GROUND TRUTH

SDE (mm)	Femur front	Femur combined	Tibia front	Tibia combined
Subject 1	1.19	2.55	1.12	1.03
Subject 2	0.62	0.66	0.87	0.72
Subject 3	0.69	0.87	1.36	1.26
Subject 4 (CT)	0.72	0.70	1.01	1.09
Subject 5	0.82	0.76	0.94	1.48
Subject 6	1.00	0.91	1.03	0.92
Subject 7	0.84	1.77	1.40	1.37
Subject 8	0.76	1.57	1.37	1.45
Subject 9	1.51	1.82	1.70	1.95
Subject 10	1.45	1.24	1.57	2.60
Mean	0.96	1.29	1.24	1.39

the lowest errors, ranging from 0.87 to 1.7 mm, with an average error of 1.24 mm. The LSO and ASM methods achieve an average of 1.84 and 2.21 mm, failing to reconstruct two and three subjects, respectively. Surface distance heatmaps for subjects #2 and #9 are given in Fig. 8, for which the best and worst reconstructions were found, respectively. A closer investigation of all ten subjects reveals that errors are mainly located in the lateral posterior condyle, as can be seen from the example. Similarly, in the case of the tibia, the largest deviations can be found on the posterior part; furthermore,

**TABLE IV**

DIRECTED MEAN SDE FROM THE MASKED RECONSTRUCTION TO THE GROUND TRUTH, AVERAGED OVER ALL TEN SUBJECTS. BEST PERFORMING ALGORITHMS ARE HIGHLIGHTED IN BOLD

SDE (mm)	Femur front	Femur combined	Tibia front	Tibia combined
GPO	<b>0.96</b>	<b>1.29</b>	<b>1.24</b>	1.39
ASM	1.72	2.05	1.84	2.18
LSO	1.57	1.6	2.21	<b>1.18</b>

the tibial plateau shows a noticeable variation. See Fig. 9 for a heatmap visualization.

When the records of the posterior parts are included, mixed results are found. All algorithms perform worse for the femur. Their rank order, however, remains the same, with the GPO being the best-performing algorithm, achieving a mean error of 1.29 mm. By contrast, performance gains can be observed for the tibia. While the ASM algorithm still fails to reconstruct several subjects, and even the error of the GPO method increases to 1.39 mm, the LSO method slightly outperforms the GPO method applied to only the frontal scan. It achieves the lowest errors of 0.71–1.66 mm, with an average of 1.18 mm.

Finally, we evaluated the overall AP and ML size as defined by [76]. The results are presented in Table V.

Regarding the computation time, all methods meet the required criteria. The CNN processes an image in 27 ms, achieving a frame rate of 37 Hz. No drops in the frame rate were observed, even with the simultaneous graphical load of the 3-D view.

The total time spent per subject lies in the range of 14 to 33 min, with 22 min spent on average. This time includes the subject getting seated, fixation of the leg, recording of two scans, repositioning and -fixation of the subject in a prone position, and recording of yet another two scans. Investigating only the time spent on recording the frontal femur, which is the largest of all four records, the time spent ranges from 62 to 142 s, with 96.4 s on average.

Exemplary results of subjects #2, #6, and #9 will be provided on request, including the ground truth model, the partial surface scan, and the full bone reconstruction.



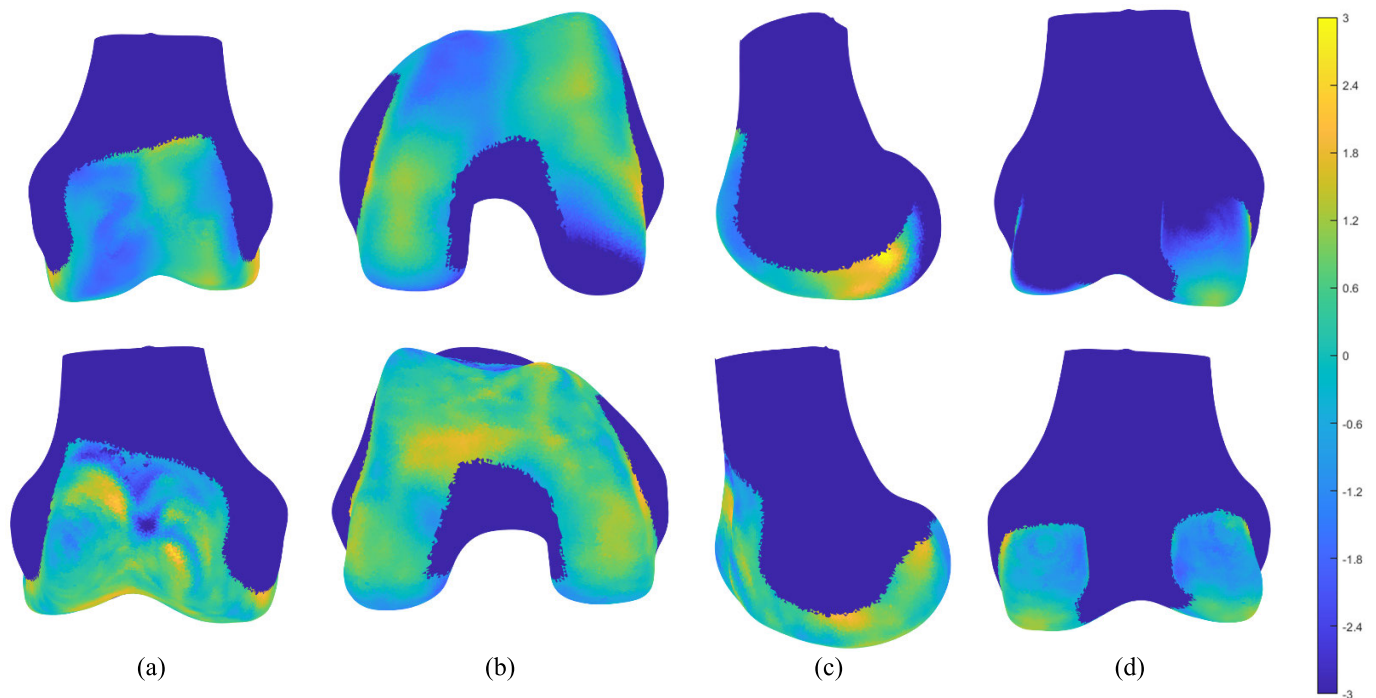


Fig. 8. Heatmap visualization of the SDE after reconstruction. From left to right: frontal, distal, medial and posterior view of the distal femur. Only the masked area is evaluated. Top row: worst reconstruction of subject #9. Note the large mismatch in the lateral condyle. Bottom row: best reconstruction of subject #2.

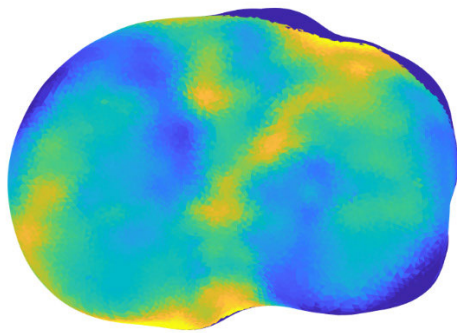


Fig. 9. Heatmap visualization of the SDE after reconstruction based on only the frontal scan for subject #6. High errors occur at the posterior and frontal crest, indicating insufficient information on the overall AP size. Similarly, the intercondylar eminence shows a rather high mismatch. Scale according to Fig. 8.

TABLE V

MEAN OF THE OVERALL AP AND ML SIZES, BASED ON EITHER MRI OR US. RESULTS FOR RECONSTRUCTIONS USING THE GPO ALGORITHM AND ONLY THE FRONTAL SCAN AS INPUT. OVERALL, A TENDENCY TO OVERSIZING CAN BE OBSERVED

(mm)	MRI	US	Difference
Femur AP	67.9	69.5	1.6
Femur ML	83.4	82	-1.4
Tibia AP	50	52.2	2.2
Tibia ML	78.2	79.9	1.7

## V. DISCUSSION

A visual comparison of the bone surface obtained from US and the reference MRI surface shows high agreement.

Neighboring bones were also, however, recorded, leading to false positive segmentation and high quantitative errors, especially in the case of the tibia, where the fibula is visible in every record. The actual surface match, investigated separately for one subject, shows an accurate surface reconstruction with errors well below 1 mm. This does not hold true when the frontal and posterior scans are combined, indicating an inadequate registration.

Investigating the accuracy after reconstruction of the full bone, an error threshold of 1 mm was achieved in the case of the femur but not of the tibia. This may be due to a large part of the tibia evaluation mask, the tibial plateau, not being imageable using US. The high errors seen in the evaluation of the original US scan could be reduced greatly, as the model is able to remove false positive segmentation of neighboring bones. In fact, even when false positive segmentation is removed manually, the reconstruction step reduces the overall error, as can be seen from the example of subject #2: the average SDE on the frontal femur scan decreases from 0.83 mm in the manually cleaned partial scan to 0.62 mm after model fitting. We hypothesize that this is due to the averaging of noise caused by segmentation errors and movement artifacts. This finding underlines that our CNN-based segmentation does not suffice for high-quality reconstruction and that a second model-fitting step is necessary.

Regarding the model-fitting method, the ASM and LSO approaches perform worse than the GPO approach on average and completely fail in some cases. The inclusion of the posterior surface scan showed only a small benefit in one evaluation and had a noticeable negative impact on all the others. As mentioned beforehand, this indicates inadequate

registration. Landmark-based registration was insufficient and ruled out early. Registering the scans on the basis of overlapping surfaces was not robust either. We, therefore, registered both scans onto the mean shape, which, of course, differs in geometry; furthermore, the information on the AP size cannot be retained by this approach. We identify this as the biggest weakness of the proposed method. Although the posterior part may be extrapolated from only the frontal scan with surprising accuracy, it remains the area of the largest surface mismatch. Apart from the difficulties when registering the posterior part, scanning the objects in a prone position also turned out to be difficult. In extension, the soft tissue covering the knee is under tension, making it hard to orient the probe orthogonally to the bone surface; furthermore, the maximum imaging depth of 7 cm was insufficient for some subjects.

Investigating the AP and ML size, a geometrical mismatch may occur on both sides (AP and ML). Accordingly, the error is expected to exceed the average SDE. The results confirm this hypothesis, especially for the tibia. Surprisingly, although no information about the posterior bone is provided to the algorithm, the AP error does not exceed the ML error. Given typical step sizes of at least 2 mm for off-the-shelf implant components, the method only based on frontal scans is sufficiently precise for implant templating in the case of the femur but not of the tibia.

Apart from the femoral ML size, the reconstructions tend to be larger than the MRI reference. One reason could be an underestimation of the average speed of sound. Salehi et al. [77] found an up to 4% mismatch in an ex vivo study, which could cause a shift up to 2.8 mm.

Comparing the errors computed for subject #4 to the average performance, we see a persistent and noticeable gap of about 0.24–0.6 mm lower errors for the CT-based evaluation. This may indicate accuracy issues in the MRI-based ground truth and an even better actual performance of our method.

In comparison to closely related previous work by Mahfouz et al. [70] and our group [54], this is the first report achieving a reconstruction of submillimetric accuracy on average; furthermore, to the best of our knowledge, it is the first small-scale in vivo evaluation. This is highly relevant, as motion artifacts may compromise accuracy, especially for larger scans. Regarding our setup, we were not able to benefit from the posterior records as their relative position to the frontal scan could not be recovered with sufficient accuracy. The main reasons for this are the low repeatability of landmark-based registration and insufficient overlap for surface-based registration. Although the setup by Mahfouz et al. [70] using electromagnetic tracking and scanning the knee in flexion could remove the need to reposition the patient, and thus, the need for registration, errors reported for electromagnetic tracking devices also have to be taken into account. Kral et al. [78] report an 0.25 mm increase in the target registration error when using an electromagnetic as opposed to an optical tracking system in a laboratory environment mimicking the operation room. Elfring [79]; however, reports up to 2.69 mm in a similar. We hypothesize that the lower reconstruction errors observed in our investigation compared to the work of Mahfouz et al. [70] were achieved by superior segmentation

and model fitting methods; however, since Mahfouz et al. [70] have not disclosed their algorithms, this must remain an assumption.

In conclusion, we proposed the first fully automatic US-based workflow that is able to achieve submillimetric in vivo reconstruction, averaged over all subjects. In future work, we will investigate the EM-based scanning procedure, as well as an automated robotic scanning approach, and quantify the individual contributing error sources to identify the potential for improvement.

## REFERENCES

- [1] OECD iLibrary. (Jan. 21, 2022). *Hip and Knee Replacement | Health at a Glance 2019 OECD Indicators | OECD iLibrary*. [Online]. Available: <https://www.oecd-ilibrary.org/sites/2fc83b9a-en/index.html?itemId=/content/component/2fc83b9a-en>
- [2] OECD. (Feb. 6, 2022). *Life Expectancy at Birth (Indicator)*. [Online]. Available: <https://data.oecd.org/healthstat/life-expectancy-at-birth.htm>
- [3] J. A. Singh, S. Yu, L. Chen, and J. D. Cleveland, "Rates of total joint replacement in the United States: Future projections to 2020–2040 using the national inpatient sample," *J. Rheumatol.*, vol. 46, no. 9, pp. 1134–1140, Sep. 2019, doi: [10.3899/jrheum.170990](https://doi.org/10.3899/jrheum.170990).
- [4] Census U.S. (Dec. 13, 2022). *Population Clock*. [Online]. Available: <https://www.census.gov/popclock/>
- [5] Y. Zhang and J. M. Jordan, "Epidemiology of osteoarthritis," *Clinics Geriatric Med.*, vol. 26, no. 3, pp. 355–369, Aug. 2010, doi: [10.1016/j.cger.2010.03.001](https://doi.org/10.1016/j.cger.2010.03.001).
- [6] M. Asseln, S. A. G. A. Grothues, and K. Radermacher, "Relationship between the form and function of implant design in total knee replacement," *J. Biomech.*, vol. 119, Apr. 2021, Art. no. 110296, doi: [10.1016/j.jbiomech.2021.110296](https://doi.org/10.1016/j.jbiomech.2021.110296).
- [7] M. A. Ritter, K. E. Davis, J. B. Meding, J. L. Pierson, M. E. Berend, and R. A. Malinzak, "The effect of alignment and BMI on failure of total knee replacement," *J. Bone Joint Surg.*, vol. 93, no. 17, pp. 1588–1596, 2011, doi: [10.2106/JBJS.J.00772](https://doi.org/10.2106/JBJS.J.00772).
- [8] M. Asseln, *Morphological and Functional Analysis of the Knee Joint for Implant Design Optimization* (Aachener Beiträge zur Medizintechnik). Düren, Germany: Shaker Verlag, 2019.
- [9] S. A. G. A. Grothues, M. Asseln, and K. Radermacher, "Variation of the femoral J-curve in the native knee," *EPiC Ser. Health Sci.*, vol. 4, pp. 86–91, Sep. 2020. [Online]. Available: <https://easychair.org/publications/open/xd3M>
- [10] S. A. G. A. Grothues and K. Radermacher, "Variation of the three-dimensional femoral J-curve in the native knee," *J. Personalized Med.*, vol. 11, no. 7, p. 592, 2021, doi: [10.3390/jpm11070592](https://doi.org/10.3390/jpm11070592).
- [11] E. J. Brinkmann and W. Fitz, "Custom total knee: Understanding the indication and process," *Arch. Orthopaedic Trauma Surg.*, vol. 141, no. 12, pp. 2205–2216, Dec. 2021, doi: [10.1007/s00402-021-04172-9](https://doi.org/10.1007/s00402-021-04172-9).
- [12] S. D. Culler, G. M. Martin, and A. Swearingen, "Comparison of adverse events rates and hospital cost between customized individually made implants and standard off-the-shelf implants for total knee arthroplasty," *Arthroplasty Today*, vol. 3, no. 4, pp. 257–263, Dec. 2017, doi: [10.1016/j.artd.2017.05.001](https://doi.org/10.1016/j.artd.2017.05.001).
- [13] V. Neginhal, W. Kurtz, and L. Schroeder, "Patient satisfaction, functional outcomes, and survivorship in patients with a customized posterior-stabilized total knee replacement," *JBJS Rev.*, vol. 8, no. 7, 2020, Art. no. e19.00104, doi: [10.2106/JBJS.RVW.19.00104](https://doi.org/10.2106/JBJS.RVW.19.00104).
- [14] K. Deep, S. Shankar, and A. Mahendra, "Computer assisted navigation in total knee and hip arthroplasty," *SICOT-J.*, vol. 3, p. 50, Jul. 2017, doi: [10.1051/sicotj/2017034](https://doi.org/10.1051/sicotj/2017034).
- [15] B. Kayani, S. Konan, A. Ayuob, E. Onochie, T. Al-Jabri, and F. S. Haddad, "Robotic technology in total knee arthroplasty: A systematic review," *EFORT Open Rev.*, vol. 4, no. 10, pp. 611–617, Oct. 2019, doi: [10.1302/2058-5241.4.190022](https://doi.org/10.1302/2058-5241.4.190022).
- [16] K. Radermacher et al., "Computer assisted orthopaedic surgery with image based individual templates," *Clin. Orthopaedics Rel. Res.*, vol. 354, pp. 28–38, Sep. 1998, doi: [10.1097/00003086-199809000-00005](https://doi.org/10.1097/00003086-199809000-00005).
- [17] L. Renson, P. Poilvache, and H. Van den Wyngaert, "Improved alignment and operating room efficiency with patient-specific instrumentation for TKA," *Knee*, vol. 21, no. 6, pp. 1216–1220, Dec. 2014, doi: [10.1016/j.knee.2014.09.008](https://doi.org/10.1016/j.knee.2014.09.008).



- [18] D. T. Mathis and M. T. Hirschmann, "Patientenspezifische instrumentierung und totalendoprothesen am knie," *Arthroskopie*, vol. 34, no. 5, pp. 342–350, Oct. 2021, doi: [10.1007/s00142-021-00460-z](https://doi.org/10.1007/s00142-021-00460-z).
- [19] M. T. Hirschmann, P. Konala, F. Amsler, F. Iranpour, N. F. Friederich, and J. P. Cobb, "The position and orientation of total knee replacement components: A comparison of conventional radiographs, transverse 2D-CT slices and 3D-CT reconstruction," *J. Bone Joint Surg.*, vol. 93, no. 5, pp. 629–633, 2011, doi: [10.1302/0301-620X.93B5.25893](https://doi.org/10.1302/0301-620X.93B5.25893).
- [20] J. R. T. Pietrzak, F. E. Rowan, B. Kayani, M. J. Donaldson, S. S. Huq, and F. S. Haddad, "Preoperative CT-based three-dimensional templating in robot-assisted total knee arthroplasty more accurately predicts implant sizes than two-dimensional templating," *J. Knee Surgery*, vol. 32, no. 7, pp. 642–648, Jul. 2019, doi: [10.1055/s-0038-1666829](https://doi.org/10.1055/s-0038-1666829).
- [21] A. Maier et al., *Medical Imaging Systems: An Introductory Guide*, Cham, Switzerland: Springer, 2018.
- [22] D. J. Brenner and E. J. Hall, "Computed tomography—An increasing source of radiation exposure," *New England J. Med.*, vol. 357, no. 22, pp. 2277–2284, Nov. 2007, doi: [10.1056/NEJMra072149](https://doi.org/10.1056/NEJMra072149).
- [23] *e-bis. Gebührenordnung für Ärzte*. Accessed: Dec. 7, 2022. [Online]. Available: <https://www.e-bis.de/goae/defaultGoae.html>
- [24] D. B. Moyano, D. A. Paraiso, and R. A. González-Lezcano, "Possible effects on health of ultrasound exposure, risk factors in the work environment and occupational safety review," *Healthcare*, vol. 10, no. 3, p. 423, Feb. 2022, doi: [10.3390/healthcare10030423](https://doi.org/10.3390/healthcare10030423).
- [25] 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, doi: [10.1016/j.ultrasmedbio.2019.12.014](https://doi.org/10.1016/j.ultrasmedbio.2019.12.014).
- [26] I. Hacihaliloglu, "3D ultrasound for orthopedic interventions," in *Intelligent Orthopaedics: Artificial Intelligence and Smart Image-Guided Technology for Orthopaedics*, G. Zheng, W. Tian, and X. Zhuang, Eds. Singapore: Springer, 2018, pp. 113–129.
- [27] A. Karamalis, "Ultrasound confidence maps and applications in medical image processing," Ph.D. dissertation, Chair Comput. Aided Med. Procedures Augmented Reality, Technische Univ. München, München, Germany, 2013.
- [28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Munich, Germany, J. Hornegger, Ed. Cham, Switzerland: Springer, Oct. 2015, pp. 234–241.
- [29] 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*. Cham, Switzerland: Springer, 2016, pp. 424–432.
- [30] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, M. J. Cardoso et al., Eds. Cham, Switzerland: Springer, 2018, pp. 3–11.
- [31] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, doi: [10.1038/s41592-020-01008-z](https://doi.org/10.1038/s41592-020-01008-z).
- [32] M. Antico et al., "Deep learning-based femoral cartilage automatic segmentation in ultrasound imaging for guidance in robotic knee arthroscopy," *Ultrasound Med. Biol.*, vol. 46, no. 2, pp. 422–435, Feb. 2020, doi: [10.1016/j.ultrasmedbio.2019.10.015](https://doi.org/10.1016/j.ultrasmedbio.2019.10.015).
- [33] A. Zaman, S. H. Park, L. Miguel, Y. Park, and S. Joung, "Real-time 3D ultrasound bone model reconstruction and its registration with MR bone model for localization of intramedullary cystic bone lesion," presented at the 15th Asian Conf. Comput. Aided Surg. (ACCAS), Tokyo, Japan, Nov. 2019.
- [34] D. Q. Duong et al., "Fully automated segmentation of alveolar bone using deep convolutional neural networks from intraoral ultrasound images," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 6632–6635, doi: [10.1109/EMBC.2019.8857060](https://doi.org/10.1109/EMBC.2019.8857060).
- [35] H. El-Hariri, K. Mulpuri, A. Hodgson, and R. Garbi, "Comparative evaluation of hand-engineered and deep-learned features for neonatal hip bone segmentation in ultrasound," in *Proc. 22nd Int. Conf. Med. Image Comput. Comput. Assist. Intervent. (MICCAI)*, Shenzhen, China, D. Shen, Ed. Cham, Switzerland: Springer, Oct. 2019, pp. 12–20.
- [36] T. Ungi et al., "Automatic spine ultrasound segmentation for scoliosis visualization and measurement," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3234–3241, Nov. 2020, doi: [10.1109/TBME.2020.2980540](https://doi.org/10.1109/TBME.2020.2980540).
- [37] S. Banerjee, S. H. Ling, J. Lyu, S. Su, and Y.-P. Zheng, "Automatic segmentation of 3D ultrasound spine curvature using convolutional neural network," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 2039–2042, doi: [10.1109/EMBC44109.2020.9175673](https://doi.org/10.1109/EMBC44109.2020.9175673).
- [38] Z. Huang et al., "Bone feature segmentation in ultrasound spine image with robustness to speckle and regular occlusion noise," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Toronto, ON, Canada, M. Elgendi and H. Shah-Mansouri, Eds. Piscataway, NJ, USA: IEEE, Oct. 2020, pp. 1566–1571.
- [39] A. Z. Alsinan, C. Rule, M. Vives, V. M. Patel, and I. Hacihaliloglu, "GAN-based realistic bone ultrasound image and label synthesis for improved segmentation," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*. Cham, Switzerland: Springer, 2020, p. 795–804.
- [40] K. C. T. Nguyen et al., "Alveolar bone segmentation in intraoral ultrasonographs with machine learning," *J. Dental Res.*, vol. 99, no. 9, pp. 1054–1061, Aug. 2020, doi: [10.1177/0022034520920593](https://doi.org/10.1177/0022034520920593).
- [41] A. Zaman, S. H. Park, H. Bang, C.-W. Park, I. Park, and S. Joung, "Generative approach for data augmentation for deep learning-based bone surface segmentation from ultrasound images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 6, pp. 931–941, Jun. 2020, doi: [10.1007/s11548-020-02192-1](https://doi.org/10.1007/s11548-020-02192-1).
- [42] C. du Toit, N. Orlando, S. Papernick, R. Dima, I. Gyacskov, and A. Fenster, "Automatic femoral articular cartilage segmentation using deep learning in three-dimensional ultrasound images of the knee," *Osteoarthritis Cartilage Open*, vol. 4, no. 3, 2022, Art. no. 100290, doi: [10.1016/j.ocarto.2022.100290](https://doi.org/10.1016/j.ocarto.2022.100290).
- [43] M. Antico et al., "Bayesian CNN for segmentation uncertainty inference on 4D ultrasound images of the femoral cartilage for guidance in robotic knee arthroscopy," *IEEE Access*, vol. 8, pp. 223961–223975, 2020, doi: [10.1109/ACCESS.2020.3044355](https://doi.org/10.1109/ACCESS.2020.3044355).
- [44] A. Kannan, A. Hodgson, K. Mulpuri, and R. Garbi, "Uncertainty estimation for assessment of 3D US scan adequacy and DDH metric reliability," in *Proc. 2nd Int. Workshop, UNSURE, 3rd Int. Workshop*, Lima, Peru, C. H. Sudre, Ed. Cham, Switzerland: Springer, Oct. 2020, pp. 97–105.
- [45] A. Kannan, A. Hodgson, K. Mulpuri, and R. Garbi, "Leveraging voxel-wise segmentation uncertainty to improve reliability in assessment of paediatric dysplasia of the hip," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 7, pp. 1121–1129, 2021.
- [46] P. Wang, M. Vives, V. M. Patel, and I. Hacihaliloglu, "Robust real-time bone surfaces segmentation from ultrasound using a local phase tensor-guided CNN," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 7, pp. 1127–1135, Jul. 2020, doi: [10.1007/s11548-020-02184-1](https://doi.org/10.1007/s11548-020-02184-1).
- [47] A. Z. Alsinan, V. M. Patel, and I. Hacihaliloglu, "Automatic segmentation of bone surfaces from ultrasound using a filter-layer-guided CNN," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 5, pp. 775–783, May 2019, doi: [10.1007/s11548-019-01934-0](https://doi.org/10.1007/s11548-019-01934-0).
- [48] S. Tang et al., "A CNN-based method to reconstruct 3-D spine surfaces from U.S. images in vivo," *Med. Image Anal.*, vol. 74, Dec. 2021, Art. no. 102221, doi: [10.1016/j.media.2021.102221](https://doi.org/10.1016/j.media.2021.102221).
- [49] S. Banerjee et al., "Light-convolution dense selection U-Net (LDS U-Net) for ultrasound lateral bony feature segmentation," *Appl. Sci.*, vol. 11, no. 21, p. 10180, Oct. 2021, doi: [10.3390/app112110180](https://doi.org/10.3390/app112110180).
- [50] K. Luan, Z. Li, and J. Li, "An efficient end-to-end CNN for segmentation of bone surfaces from ultrasound," *Comput. Med. Imag. Graph.*, vol. 84, Sep. 2020, Art. no. 101766, doi: [10.1016/j.compmedimag.2020.101766](https://doi.org/10.1016/j.compmedimag.2020.101766).
- [51] A. Rahman, W. G. C. Bandara, J. M. J. Valanarasu, I. Hacihaliloglu, and V. M. Patel, "Orientation-guided graph convolutional network for bone surface segmentation," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*. Cham, Switzerland: Springer, 2022.
- [52] M. Duhnhofer et al., "Siam-U-Net: Encoder-decoder Siamese network for knee cartilage tracking in ultrasound images," *Med. Image Anal.*, vol. 60, Feb. 2020, Art. no. 101631, doi: [10.1016/j.media.2019.101631](https://doi.org/10.1016/j.media.2019.101631).
- [53] G. Kompella et al., "Segmentation of femoral cartilage from knee ultrasound images using mask R-CNN," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 966–969, doi: [10.1109/embc.2019.8857645](https://doi.org/10.1109/embc.2019.8857645).
- [54] B. Hohlmann and K. Radermacher, "Augmented active shape model search—Towards 3D ultrasound-based bone surface reconstruction," in *Proc. EPiC Ser. Health Sci.*, 2020, pp. 117–121.

- [55] P. Broessner, B. Hohlmann, and K. Radermacher, "Ultrasound-based navigation of scaphoid fracture surgery," in *Proc. German Workshop Med. Image Comput.*, Heidelberg, Germany, K. H. Maier-Hein, T. M. Deserno, H. Handels, A. Maier, C. Palm, and T. Tolxdorff, Eds. Wiesbaden, Germany: Springer Vieweg, Jun. 2022, pp. 28–33.
- [56] B. Hohlmann, P. Brößner, K. Welle, and K. Radermacher, "Segmentation of the scaphoid bone in ultrasound images," *Current Directions Biomed. Eng.*, vol. 7, no. 1, pp. 76–80, Aug. 2021, doi: [10.1515/cdbme-2021-1017](https://doi.org/10.1515/cdbme-2021-1017).
- [57] B. Hohlmann, J. Glanz, and K. Radermacher, "Segmentation of the distal femur in ultrasound images," *Current Directions Biomed. Eng.*, vol. 6, no. 1, Sep. 2020, Art. no. 20200034, doi: [10.1515/cdbme-2020-0034](https://doi.org/10.1515/cdbme-2020-0034).
- [58] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.
- [59] Y. Zhou, A. Rakkunedeth, C. Keen, J. Knight, and J. L. Jaremko, "Wrist ultrasound segmentation by deep learning," in *Proc. 20th Int. Conf. Artif. Intell. Med. AIME*, Halifax, NS, Canada, 1st ed., M. Michalowski, S. S. R. Abidi, and S. Abidi, Eds. Cham, Switzerland: Springer, Jun. 2022, pp. 230–237.
- [60] A. Z. Alsinan, V. M. Patel, and I. Hacihaliloglu, "Bone shadow segmentation from ultrasound data for orthopedic surgery using GAN," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 9, pp. 1477–1485, Sep. 2020, doi: [10.1007/s11548-020-02221-z](https://doi.org/10.1007/s11548-020-02221-z).
- [61] K. Morooka, M. Nakamoto, and Y. Sato, "A survey on statistical modeling and machine learning approaches to computer assisted medical intervention: Intraoperative anatomy modeling and optimization of interventional procedures," *IEICE Trans. Inf. Syst.*, vol. E96.D, no. 4, pp. 784–797, 2013.
- [62] N. Sarkalkan, H. Weinans, and A. A. Zadpoor, "Statistical shape and appearance models of bones," *Bone*, vol. 60, pp. 129–140, Mar. 2014, doi: [10.1016/j.bone.2013.12.006](https://doi.org/10.1016/j.bone.2013.12.006).
- [63] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Med. Image Anal.*, vol. 13, no. 4, pp. 543–563, Aug. 2009, doi: [10.1016/j.media.2009.05.004](https://doi.org/10.1016/j.media.2009.05.004).
- [64] I. Hacihaliloglu, A. Rasoulian, R. N. Rohling, and P. Abolmaesumi, "Statistical shape model to 3D ultrasound registration for spine interventions using enhanced local phase features," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 16. Cham, Switzerland: Springer, 2013, pp. 361–368, doi: [10.1007/978-3-642-40763-5\\_45](https://doi.org/10.1007/978-3-642-40763-5_45).
- [65] I. Hacihaliloglu, A. Rasoulian, R. N. Rohling, and P. Abolmaesumi, "Local phase tensor features for 3-D ultrasound to statistical shape+pose spine model registration," *IEEE Trans. Med. Imag.*, vol. 33, no. 11, pp. 2167–2179, Nov. 2014, doi: [10.1109/TMI.2014.2332571](https://doi.org/10.1109/TMI.2014.2332571).
- [66] A. Rasoulian, R. N. Rohling, and P. Abolmaesumi, "Augmentation of paramedian 3D ultrasound images of the spine," in *Proc. 4th Int. Conf. Inf. Process. Comput.-Assist. Intervent. (IPCAI)*, Heidelberg, Germany, D. Barratt, S. Cotin, G. Fichtinger, P. Jannin, and N. Navab, Eds. Berlin, Germany: Springer, Jun. 2013, pp. 51–60.
- [67] E. M. A. Anas et al., "Registration of a statistical model to intraoperative ultrasound for scaphoid screw fixation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 6, pp. 957–965, Jun. 2016, doi: [10.1007/s11548-016-1370-y](https://doi.org/10.1007/s11548-016-1370-y).
- [68] C. Hänisch, J. Hsu, E. Noorman, and K. Radermacher, "Model based reconstruction of the bony knee anatomy from 3D ultrasound images," in *Proc. 15th Annu. Meeting Int. Soc. Comput. Assist. Orthopaedic Surg.*, Vancouver, BC, Canada, Jun. 2015, pp. 1–6.
- [69] B. Hohlmann and K. Radermacher, "The interleaved partial active shape model (IPASM) search algorithm—Towards 3D ultrasound-based bone surface reconstruction," in *Proc. 19th Annu. Meeting Int. Soc. Comput. Assist. Orthopaedic Surg.*, 2019, pp. 172–177, doi: [10.29007/rbgl](https://doi.org/10.29007/rbgl).
- [70] M. R. Mahfouz, E. E. A. Fatah, J. M. Johnson, and R. D. Komistek, "A novel approach to 3D bone creation in minutes 3D ultrasound," *Bone Joint J.*, vol. 103-B, no. 6, pp. 81–86, 2021, doi: [10.1302/0301-620X.103B6.BJJ-2020-2455.R1](https://doi.org/10.1302/0301-620X.103B6.BJJ-2020-2455.R1).
- [71] A. Fedorov et al., "3D slicer as an image computing platform for the quantitative imaging network," *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1323–1341, Nov. 2012, doi: [10.1016/j.mri.2012.05.001](https://doi.org/10.1016/j.mri.2012.05.001).
- [72] 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*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 833–851.
- [73] P. Brößner, B. Hohlmann, and K. Radermacher, "Transformer vs. CNN—A comparison on knee segmentation in ultrasound images," in *Proc. 22nd Annu. Meeting Int. Soc. Comput. Assist. Orthopaedic Surg.*, 2022, pp. 24–31.
- [74] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—Their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995, doi: [10.1006/cviu.1995.1004](https://doi.org/10.1006/cviu.1995.1004).
- [75] V. Blanz, A. Mehl, T. Vetter, and H.-P. Seidel, "A statistical method for robust 3D surface reconstruction from sparse data," in *Proc. 2nd Int. Symp. 3D Data Process., Vis. Transmiss. (DPVT)*, 2004, pp. 293–300.
- [76] M. Mahfouz, E. E. A. Fatah, L. S. Bowers, and G. Scuderi, "Three-dimensional morphology of the knee reveals ethnic differences," *Clin. Orthopaedics Rel. Res.*, vol. 470, no. 1, pp. 172–185, 2012, doi: [10.1007/s11999-011-2089-2](https://doi.org/10.1007/s11999-011-2089-2).
- [77] M. Salehi, R. Prevost, J.-L. Moctezuma, N. Navab, and W. Wein, "Precise ultrasound bone registration with learning-based segmentation and speed of sound calibration," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2017, pp. 682–690.
- [78] F. Kral, E. J. Puschban, H. Reichelmann, F. Pedross, and W. Freysinger, "Optical and electromagnetic tracking for navigated surgery of the sinuses and frontal skull base," *Rhinol. J.*, vol. 49, no. 3, pp. 364–368, Aug. 2011, doi: [10.4193/Rhino10.177](https://doi.org/10.4193/Rhino10.177).
- [79] R. Elfring, "Störungskompensation und Optimierung des elektromagnetischen Trackings in der computerunterstützten Chirurgie," Ph.D. dissertation, Aachener Beiträge zur Medizintechnik, Shaker Verlag, Düren, Germany, 2012.