Computers in Biology and Medicine 178 (2024) 108794

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Automatic Segmentation and Alignment of Uterine Shapes from 3D Ultrasound Data

Eva Boneš [a,*], Marco Gergolet [b], Ciril Bohak [a,c], Žiga Lesar [a], Matija Marolt [a]

[a] University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, Ljubljana, 1000, Slovenia
[b] University of Ljubljana, Faculty of Medicine, Vrazov trg 2, Ljubljana, 1000, Slovenia
[c] King Abdullah University of Science and Technology, Visual Computing Center, Thuwal, 23955-6900, Saudi Arabia

ARTICLE INFO

Keywords:
Uterus segmentation
Volumetric ultrasound
3D alignment

ABSTRACT

Background: The uterus is the most important organ in the female reproductive system. Its shape plays a critical role in fertility and pregnancy outcomes. Advances in medical imaging, such as 3D ultrasound, have significantly improved the exploration of the female genital tract, thereby enhancing gynecological healthcare. Despite well-documented data for organs like the liver and heart, large-scale studies on the uterus are lacking. Existing classifications, such as VCUAM and ESHRE/ESGE, provide different definitions for normal uterine shapes but are not based on real-world measurements. Moreover, the lack of comprehensive datasets significantly hinders research in this area. Our research, part of the larger NURSE study, aims to fill this gap by establishing the shape of a normal uterus using real-world 3D vaginal ultrasound scans. This will facilitate research into uterine shape abnormalities associated with infertility and recurrent miscarriages.

Methods: We developed an automated system for the segmentation and alignment of uterine shapes from 3D ultrasound data, which consists of two steps: automatic segmentation of the uteri in 3D ultrasound scans using deep learning techniques, and alignment of the resulting shapes with standard geometrical approaches, enabling the extraction of the normal shape for future analysis. The system was trained and validated on a comprehensive dataset of 3D ultrasound images from multiple medical centers. Its performance was evaluated by comparing the automated results with manual annotations provided by expert clinicians.

Results: The presented approach demonstrated high accuracy in segmenting and aligning uterine shapes from 3D ultrasound data. The segmentation achieved an average Dice similarity coefficient (DSC) of 0.90. Our method for aligning uterine shapes showed minimal translation and rotation errors compared to traditional methods, with the preliminary average shape exhibiting characteristics consistent with expert findings of a normal uterus.

Conclusion: We have presented an approach to automatically segment and align uterine shapes from 3D ultrasound data. We trained a deep learning nnU-Net model that achieved high accuracy and proposed an alignment method using a combination of standard geometrical techniques. Additionally, we have created a publicly available dataset of 3D transvaginal ultrasound volumes with manual annotations of uterine cavities to support further research and development in this field. The dataset and the trained models are available at https://github.com/UL-FRI-LGM/UterUS.

1. Introduction

The uterus plays a very important role in women's reproductive outcomes. Recent advances in medical imaging and technology (*e.g.*, 3D ultrasound) improved the exploration of the female genital tract and, with it, increased the quality of gynecological health care.

While the average size and distribution data for most organs, such as the liver, brain, and heart, are well-established [1–3], there is no large-scale study of the uterus that examines its normal characteristics. Some uterine shape classifications, such as Vagina Cervix Uterus Adnex-associated Malformation (VCUAM) [4] and European Society of Human Reproduction and Embryology (ESHRE)/European Society for Gynaecological Endoscopy (ESGE) [5], do include separate classes for the normal uterus. While VCUAM classification did not define either the measures or the shape of the normal uterus, ESHRE/ESGE classification

* Corresponding author.
E-mail addresses: eva.bones@lgm.fri.uni-lj.si (E. Boneš), marco.gergolet@gmail.com (M. Gergolet), ciril.bohak@fri.uni-lj.si (C. Bohak), ziga.lesar@fri.uni-lj.si (Ž. Lesar), matija.marolt@fri.uni-lj.si (M. Marolt).

https://doi.org/10.1016/j.compbiomed.2024.108794
Received 31 December 2023; Received in revised form 18 June 2024; Accepted 19 June 2024
Available online 27 June 2024

defined the normal uterus as an organ having "either straight or curved interostial line but with an internal indentation at the fundal midline not exceeding 50% of the uterine wall thickness", avoiding the use of absolute values.

Our research is part of a larger study called NURSE (**N**ormal **U**te**R**ine as**SE**ssment) that aims to fill the gap by finding measurements that define a normal uterus, encompassing various lengths, thicknesses, and angles of the uterine wall. With a new standard that would clearly define the normal shape, research into possible shape abnormalities in women with unexplained infertility or repeated miscarriages could become feasible and become part of routine diagnostic procedures for primary or secondary infertility. The study is being conducted in collaboration with several medical centers from different European countries with a large number of female subjects.

Professionals have traditionally used the Virtual Organ Computer-Aided Analysis (VOCAL) system to make uterine measurements by defining a cross-section of the 3D volume on 2D images. However, this method inherently limits accuracy due to its reliance on 2D views and is prone to human error and inter-operator variability. Moreover, the VOCAL system necessitates drawing the uterine cavity's contour in multiple projections, requiring the process to be repeated 30 times with a 12-degree rotation of the image, as detailed in the GE User Manual. Unfortunately, this often results in disappointing outcomes, such as spikes or unnatural distortions that fail to accurately represent the true shape of the uterine cavity [6].

To overcome these limitations, we propose a method that extracts and aligns 3D uterine shapes from 3D ultrasound volumes. This approach will allow us to create a reference model of a normal 3D uterus that will, in combination with the normal shape assessed statistically from 2D measurements by the NURSE study, aid infertility specialists in comparing the uterine cavity shape of an individual patient with that of the established normal model, enhancing diagnostic and treatment accuracy. 3D views obtained from aligned 3D shapes can provide insights that are unattainable through manual 2D measurements and reduce the scope for human error. Automating the entire process not only minimizes human involvement but also significantly speeds up the processing of large data volumes. Moreover, the resulting 3D shapes facilitate easy visualization, serving as a valuable educational tool.

The main contributions of our paper are:

1. We provide a trained neural model for **automatic segmentation** of uterine cavities in 3D ultrasound data. The model is based on the nnU-net architecture [7] and gives good results on 3D ultrasound data acquired in various hospitals by different operators.
2. We present a novel method for **automatic alignment** of extracted 3D uterine shapes. The proposed method is based on geometric and mathematical approaches and can be used to align and compare the 3D uterine shapes of different patients.
3. We present a **publicly available dataset** of 3D transvaginal ultrasound volumes with manual annotations of uterine cavities. The dataset will help expand the currently limited pool of freely available volumetric ultrasound data. It will enable the development of new segmentation algorithms, facilitate comparative studies in this area, and could also be used to solve other problems in the broader field of gynecology or medical image analysis.

1.1. Background

The uterus is a hollow, pear-shaped muscular (smooth muscle) structure. Anatomically, the uterus consists of three main parts: the uterine corpus or body, housing the uterine cavity; the isthmus — the part where the uterine cavity starts to narrow; and the uterine cervix. Fig. 1 presents the rough anatomy of the uterus. In our research, we focused mainly on the cavity — therefore, unless stated otherwise, we will use the term *uterus* to refer to the uterine cavity. While most of us can easily distinguish the rough anatomy of the uterus in a coronal view (frontal view of the human body) as depicted in Fig. 1, it is crucial to recognize the uterus as a three-dimensional object situated in a three-dimensional environment, often appearing curved in the lateral view (side view of the human body). The lateral view and the position of the uterus inside the body are presented in Fig. 2. This curvature, among other complexities, poses a challenge for the comparison and automatic alignment of uteri.

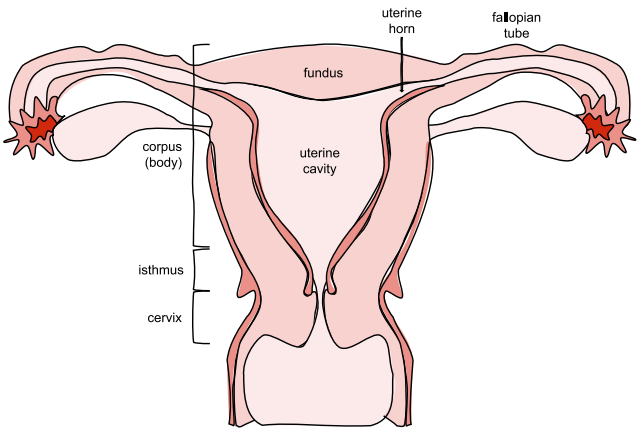


Fig. 1. Rough anatomy of the uterus. Image inspired by Nott [8].

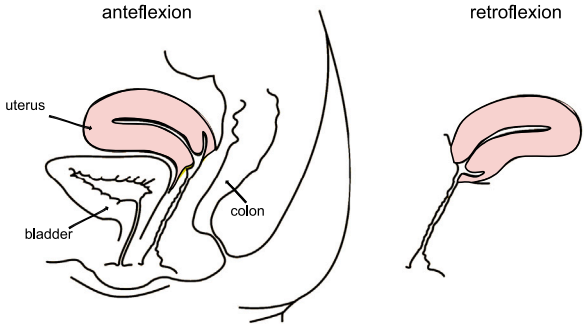


Fig. 2. Position of the uterus in the body. Image adapted from Ellis [9].

Ultrasound, a widely used medical imaging technique, offers valuable insights into soft tissues and organs due to its minimal invasiveness, portability, and cost-effectiveness. However, challenges remain in ultrasound analysis, particularly in the presence of noise and artifacts. Speckle noise, resulting from the scattering of ultrasound waves, can introduce random fluctuations in the received signal, complicating image analysis. Moreover, various artifacts such as acoustic shadowing, reverberation, mirror image artifacts, and acoustic enhancement, can hinder accurate diagnostics [10–13]. Despite the noise, medical professionals are trained and skilled at telling apart artifacts from true anatomical structures, but it is challenging for the automated approaches to do the same [14].

3D ultrasound volumes are generated by capturing a sequence of 2D images using specialized 3D ultrasound probes, which can adjust the direction of the ultrasound beam either manually or automatically. These probes are often equipped with an electromagnetic position sensor or an electrical gyroscope to ensure accurate spatial orientation of each image captured. The computer system records these images and gathers the necessary spatial data to construct the 3D volume. The quality of these 3D ultrasound volumes heavily relies on the accurate placement of the ultrasound probe and the expertise of the medical professional performing the examination. Unlike many other imaging modalities, where the patient's position and the imaging direction

are predetermined, ultrasound examinations require the operator to skillfully manipulate the probe to obtain optimal images and, thus, optimally constructed volumes. This operator-dependent nature of ultrasound poses unique challenges, especially in achieving consistent quality across different examinations and among various operators.

Moreover, ultrasound does not provide uniform resolution across all axes. The imaging plane that is displayed on the ultrasound screen following volume acquisition, known as the acquisition plane, offers the best visualization. It is critical that the structure of interest, in our case, the uterus, is well-represented in this initial plane to ensure the accuracy of the final results [15].

1.2. Related work

1.2.1. Ultrasound segmentation

As discussed in Section 1.1, ultrasound images present significant challenges for automated analysis. From noise and artifacts resulting from the physical properties of ultrasound propagation to the fact that the quality of the images is highly dependent on the operator performing the ultrasound. Recently, however, technological advancements in ultrasound equipment, such as improved ultrasound probes, higher resolution, new digital systems, etc., have substantially enhanced image quality [16]. Furthermore, the advent of deep learning has revolutionized data analysis in this field, particularly in noise reduction and feature extraction, making it a highly active area of research.

The goals of biomedical data analysis can be divided into several areas: classification, segmentation, detection, and reconstruction. Recently, most analysis approaches have been based on deep neural networks. Authors use various architectures, from convolutional neural networks (CNN) effectively applied for analyzing different organs such as the liver [17], breast [18], thyroid [19], and arteries [20] to recurrent neural networks (RNN) [21,22], encoder–decoder networks [23,24], and generative adversarial networks (GAN) [25]. Fully convolutional network (FCN) [26] based approaches proved successful, mostly when stacking multiple simpler FCN models like the methods proposed by Zhang et al. [27] and Wu et al. [28]. Mishra et al. [29] added deep supervision [30] to overcome the difficulty of training FCNs on complex ultrasound data as it generally needs more layers that are not possible with FCN alone. The U-Net architecture [31] as an extension of FCN was explicitly designed for automatic semantic segmentation of biomedical data and has been very successful. Several extensions to the U-Net architecture have been proposed. Valanarasu et al. [32] combined U-Net with the Ki-Net encoder–decoder architecture to improve the segmentation of objects with blurry and noisy edges. Jain et al. [33] fused U-Net(+) and SegNet to address the placenta segmentation in ultrasound, where they use data similar to our work. Kumar et al. [34] used a Multi U-Net model to perform real-time segmentation of breast ultrasound structures. Xu et al. [35] recently introduced the Vector Self-Attention Layer (VSAL) for ultrasound data, improving object detection in various aspects while maintaining equivariance properties. The nnU-net model [7] as an extension to U-Net has proven successful in the broader field of biomedical data. Despite being a general model, it achieved top results in 33 of 53 tasks that are commonly used to evaluate segmentation models and came close to the best methods in the remaining tasks.

Attention mechanisms have further improved segmentation performance by allowing models to focus on the most relevant parts of the input, enhancing their capability to handle complex biomedical images. These mechanisms have been integrated into various neural network architectures. For instance, Oktay et al. [36] proposed the Attention U-Net, which integrates attention gates into the U-Net architecture to focus on relevant regions. Similarly, Wang et al. [37] introduced a residual attention network that combines residual learning with attention mechanisms to enhance feature representation, which has later been used to improve segmentation performance in challenging medical image datasets [38]. Hatamizadeh et al. [39] proposed Swin-UNETR, which incorporates Swin Transformers [40] into a U-Net-based architecture, leveraging self-attention mechanisms to capture long-range dependencies in medical images, surpassing the state-of-the-art in various biomedical segmentation tasks at the time of publication.

Despite extensive research and progress in the field, the problem of automatic segmentation of the uterus in volumetric ultrasound data remains challenging. In a recently presented approach to 3D segmentation, Behboodi et al. [41] used a 2D approach where the model was run on each slice in all three dimensions, and results were merged to produce a 3D segmentation. This method loses some spatial information that can be valuable for segmenting volumetric data. The method mainly has problems with accurately labeling edges but is relatively successful in segmenting intermediate slices. Therefore, the approach is more suitable for assisting in manual segmentation rather than replacing the entire process. Shahedi et al. [42] also proposed a semi-automatic approach to 3D segmentation, using 3D U-Net, where an operator was employed to select the bounding box of the uterus and placenta.

Kong et al. [43] introduced an approach for fetal and uterine segmentation in volumetric ultrasound data using a fully convolutional neural network (FCN) and bidirectional long short-term memory (BiLSTM). The encoder part of the FCN extracts relevant features from 2D slices, while the BiLSTM layer is responsible for finding correlations between slices. Two branches of the FCN decoder produce the final segmentation of the fetus and uterus. In their research, the data used differs significantly from our dataset as it was acquired at later stages of pregnancy when the uterus has a different shape and appearance. Additionally, their data was obtained using abdominal ultrasound, whereas ours was acquired using transvaginal ultrasound.

Another modality for uterus diagnostics is magnetic resonance imaging (MRI) – Kurata et al. [44] presented an approach to 2D segmentation of uteri, showing that the segmentation accuracy does not reduce significantly with the presence of disorders. They achieved a Dice coefficient (DSC) [45,46] of around 0.8. More recently, Zhou et al. [47] presented a study on 3D reconstruction and automatic segmentation of the uterine cavity. They used the recurrent saliency transformation network (RSTN) [48] and achieved similar scores. The MRI data, however, typically provides higher resolution and less noisy images than ultrasound, making segmentation easier.

1.2.2. 3D object alignment

3D data alignment is a well-researched topic with applications spanning numerous fields, including robotics, 3D scanning, and medical image registration. Classical point-based approaches, such as Iterative Closest Point (ICP) [49,50], Coherent Point Drift (CPD) [51], and Principal Component Analysis (PCA) [52], are commonly employed for this purpose. Additionally, geometric properties, like the convex hull [53], provide alternative methods for alignment.

Feature-based methods, which identify and link salient points on 3D models – such as corners, edges, and key points – are also prevalent. Notable examples include the Scale-Invariant Feature Transform (SIFT) [54], Speeded-Up Robust Features (SURF) [55], and Random Sample Consensus (RANSAC) [56].

Within the medical context, Statistical Shape Models (SSMs) [57] and amongst them, Active Shape models [58] and Active Appearance models [59] are often utilized due to their ability to align well-defined and consistent shapes across similar datasets. These models build upon the analysis of shape variations within a population, making them particularly useful in scenarios where anatomical landmarks are consistent and well-defined. However, their efficacy diminishes when applied to highly variable anatomical structures.

These traditional methods generally assume that the objects to be aligned exhibit a high degree of similarity and can be accurately described by a set of predetermined landmarks. This assumption often proves problematic in medical settings, where anatomical variations

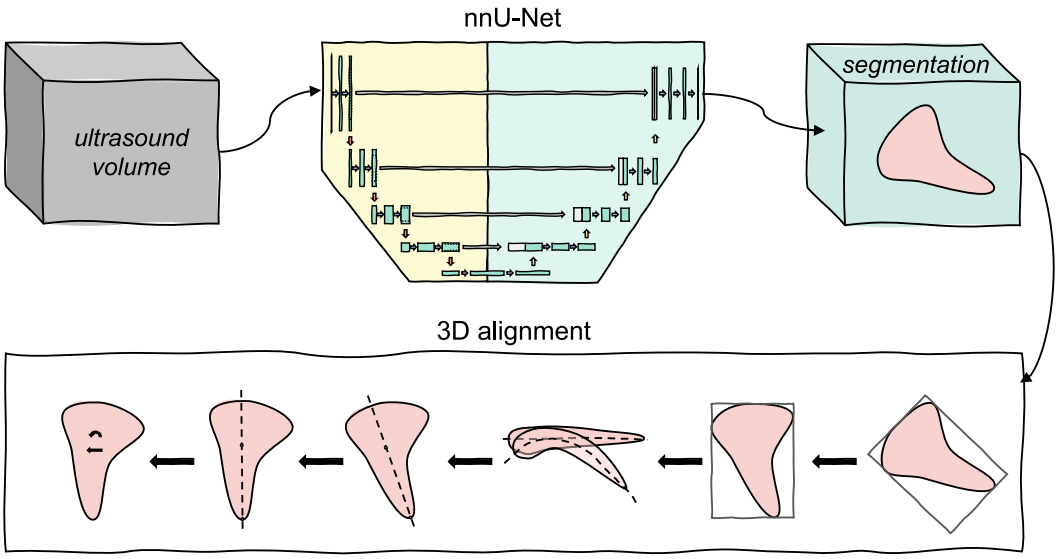


Fig. 3. Diagram of the procedure from ultrasound volume to the final result - aligned uteri.

from patient to patient present unique alignment challenges — especially with organs like the uterus, which not only vary between individuals but also undergo physiological changes.

To overcome this, deep learning methods [60] have gained importance in this area, as they have the ability to align objects with significantly different shapes, given a sufficient amount of data. The recently introduced ART [61] is a neural network module that can be integrated into various architectures for 3D tasks (such as point cloud autoencoding, shape alignment, shape interpolation, human pose transfer, etc.) as the authors have shown that having aligned shapes improves their performance. The module learns to rotate the input shapes to their canonical orientation through self-supervised learning, eliminating the need for gold standards. Although the approach aims to find the orientation for each object individually and was not created solely for object alignment, it implicitly aligns objects by moving them into the same coordinate system. The authors have shown that even for objects that are not identical (*e.g.*, different car shapes, different dog breeds), the network finds the same canonical orientations and aligns the objects with each other. However, a significant amount of data is required to train the model successfully.

Despite the advancements and promising results shown by deep learning methods, our decision to adopt a geometrical approach for 3D uterine alignment was driven by several factors. Firstly, the high variability in uterine shapes presents a significant challenge for training-based methods, which typically require large, annotated datasets to achieve reliable performance. In the context of our study, such extensive datasets are not readily available. Secondly, the anatomical variability of the uterus, both between different individuals and within the same individual over time, complicates the application of standard alignment techniques. While ICP and CPD are effective for aligning objects with high similarity, they struggle with significant anatomical variability and lack of consistent landmarks in uterine shapes. Geometrical methods, on the other hand, do not rely on large training datasets and are inherently more adaptable to variability. Additionally, these methods allow for more interpretable and controllable alignment processes, which is crucial for ensuring clinical relevance and accuracy. By focusing on a purely geometrical approach, we aim to provide a robust, efficient, and transparent solution for aligning 3D uterine shapes, thereby facilitating more accurate and consistent analyses in medical research and practice (see Fig. 3).

2. Materials and methods

In this section, we describe the main contributions of the paper. We collected a dataset of transvaginal 3D ultrasound scans (the UterUS dataset) and annotated the uterine cavities as described in Section 2.1. To enable visualization and comparison of such scans, we developed a procedure that includes two key steps: segmentation (described in Section 2.2) and alignment (described in Section 2.3).

2.1. The uterUS dataset

In collaboration with several medical centers across Europe, we have collected a set of 298 3D ultrasound volumes of the uterus. The data was collected from female subjects, categorized into three groups:

1. women who have never been pregnant aged between 18 and 28 in order to avoid acquired anomalies of the uterus that have the tendency to increase with age — *general population* (G),
2. women with at least 6 months of *unexplained* causes of infertility (normal semen analysis, tubal patency, no ovulation disorders) (I),
3. women with 2 or more spontaneous miscarriages (normal karyotype, no antiphospholipid syndrome or thrombophilia) (M).

The exclusion criteria for participating in the study were as follows:

1. women who have been pregnant (since pregnancy may modify several measurements of the uterus),
2. women with acquired uterine abnormalities, such as myomas and adenomyosis,
3. women with endometrial polyps (who could be included after polypectomy),
4. women with a unicornuate uterus.

In all subjects included in the study, examinations were conducted during the proliferative phase of the menstruation cycle. This choice was made because, during the secretory phase, a thickened endometrium can create irregularities in the uterine cavity, as shown by Saravelos et al. [62]. The ultrasound exam was done between the 11th and 14th days of the menstrual cycle when the lining of the uterus is clearly visible. The volume acquisition was performed in a sagittal projection with the region of interest (ROI) maximized. To minimize artifacts, patients were instructed to hold their breath for a few seconds during the acquisition process. One volume was acquired per patient. At the current stage, the study focused on determining the shape and

measurements of the general population. We have 141 volumes from group G, 96 volumes from group I, and 9 volumes from group M. For the remaining samples, we do not yet have metadata. In the future, we plan to extend the dataset to include more samples from other classes for further analysis.

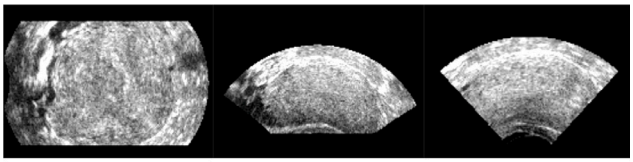
Most of our collected volumes were acquired with General Electric (GE) ultrasound equipment. The acquired ultrasound data were originally in a toroidal system, as used by the GE ultrasound machines [63]. One dimension represents radial distance, while the other two dimensions represent angles, and the voxel spacing in this system is not isotropic. To facilitate analysis and visualization, we transformed this data into a Cartesian coordinate system using the 3D Slicer [64] extension SlicerHeart [65]. During the transformation to Cartesian coordinates, the inherently cone-shaped ultrasound data is enclosed within a cuboid. Any unoccupied voxels within this cuboid are set to 0. This approach preserves the anatomical shapes of the organs and avoids distorting them, even though it results in some voxel space being unused (black voxels on the borders). This also proved to be more understandable when visualized by both the annotator working on the uterine annotations and the scientists responsible for verifying the accuracy of the annotations, as it is consistent with the 2D representations they are used to.

The sizes of the volumes range from $135 \times 109 \times 71$ to $336 \times 267 \times 169$ voxels, with brightness, contrast, and noise levels varying within the dataset. The voxel spacing in all volumes is isotropic, with a uniform spacing of $0.50\,\text{mm}$ in all directions. We also acquired 44 volumes with Samsung ultrasound machines, making the dataset multi-device. The volumes were converted from their proprietary format to the Neuroimaging Informatics Technology Initiative (NIfTI) format [66] with a tool developed by Samsung's R&D team.

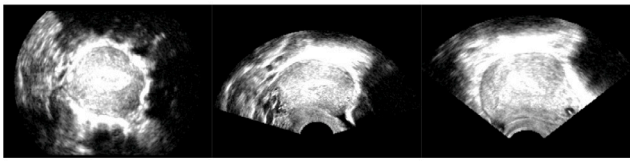
For training and evaluation of segmentation models, we manually annotated the uterine cavity in the acquired volumes. For annotation, we used 3D Slicer, an open-source software for visualization, analysis, and annotation of medical, biomedical, and other 3D data. Volumes imported into 3D Slicer can be viewed in slices along axial, coronal, and sagittal planes. Annotations were made by coloring the uterine cavity in each slice along a selected axis.

The annotation procedure proved challenging as accurate identification of the uterus is sometimes difficult because the quality of ultrasound volumes can vary significantly; in some volumes, the uterine region may be clearly visible, whereas in others, the boundaries are barely discernible. The quality of the data depends on several factors. It may be influenced by the ultrasound machine's quality and the technician's skill, as discussed in Section 1. Since we collected data from different hospitals, all these factors also play a role in our data set. Because ultrasound travels through tissue, the patient's anatomy and the tissue's composition around the region of interest also play a role [67]. Unforeseen hardware maintenance issues, such as faulty transducer cable shielding, an unconnected cable in the room that may act as an antenna, excessive dust, and oxidation in the system, etc., can also add noise [68]. Due to these factors, the volumes may have different noise, contrast, and brightness levels, all affecting the overall quality of the volumes, as shown in Fig. 4. For this reason, annotating along a single axis was rarely sufficient, and in most cases, information from all three dimensions had to be used for accurate annotations. The final annotations were obtained by combining information from all views and adjusting the annotations over each dimension. The 3D shape was checked regularly during the annotation process to ensure that the results were correct. As the final step, minimal smoothing of the segmentation was employed to ensure continuity between slices.

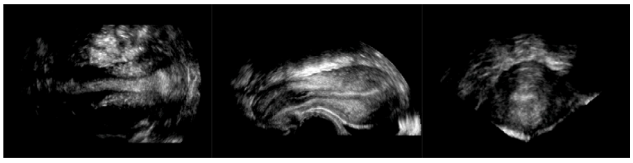
The most challenging part of the annotation process was delineating the cervix. The length and appearance of the cervix vary significantly across ultrasound volumes due to differences in ultrasound diagnostics and patient anatomy. In addition, the visibility of the cervix decreases toward its end (closer to the vagina) because ultrasound quality is poorer at greater distances from the region of interest (the uterine cavity). Even gynecologists measuring the distances manually did not always agree on where the cervix ends. However, the slightly less accurate cervix annotations did not significantly affect the overall results of the trained segmentation models.



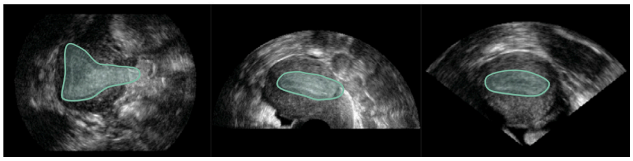
(a) A volume with high level of noise.



(b) A very bright volume.



(c) A volume with low contrast and brightness.



(d) A volume of good quality with distinct boundaries of the uterine cavity. The borders shown are not precise, but are here to help the reader determine where the uterus is positioned.

Fig. 4. Slices of volumes on different planes with different levels of quality.

The entire dataset was annotated by a single annotator in order to minimize variability. The annotator, while not a medical professional, received training from a gynecologist to accurately identify and mark the boundaries of the uterus. In addition, the final annotations were reviewed by a gynecologist. Annotating a volume took about 20 min on average, although the time varied depending on the volume's quality and the uterus's shape. Some volumes required extensive adjustments and adaptations along one axis after initial segmentation on another. An example of an annotated volume is shown in Fig. 5.

Volumes with insufficiently visible uterine regions were not annotated because poorly annotated data would adversely affect further steps. In addition, data obtained later in the research was not yet manually annotated but is included in the publicly available dataset. The dataset, named **UterUS** (**Uter**ine **U**ltra**S**ound), thus contains 141 3D ultrasound volumes with manual annotations of the uterine region and 113 unannotated volumes from 11 different hospitals. The volumes and their associated segmentations are stored in the NiFTi file format. The dataset is available at https://github.com/UL-FRI-LGM/UterUS under the CC-BY-NC-SA 4.0 licence.[1]

2.2. Automatic segmentation

Our goal was to develop a method for automatically segmenting uteri in 3D ultrasound data. Given the rapid advances in the

[1] https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode

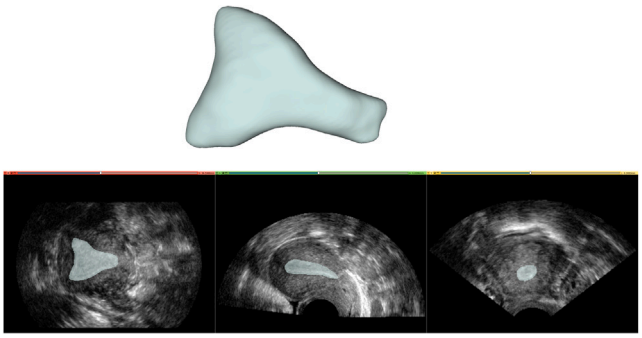


Fig. 5. An annotated volume visualized with the 3D Slicer tool. Above is the 3D representation of the annotated region, and below are 2D projections that allow movement along the coronal (red), axial (green), and sagittal (yellow) planes.

field and the frequent introduction of new machine-learning architectures in various domains, we began exploring the broader biomedical data segmentation field. This included segmentation for different data types, such as microscopic, X-ray, MRI, and CT volumes. On a subset of the UterUS data, we first evaluated several architectures, including HighRes3DZMUNet [69], an extension of the U-Net architecture; Med3D [70], which is pre-trained on various medical data; contrastive learning [71], which incorporates unsupervised learning and allows for the use of unannotated data; nnU-net [7], an extension of the U-Net [31]; and Swin UNETR [39], combining the characteristics of Swin transformers [40] and U-Net. The best results were obtained with nnU-net and Swin UNETR, which we chose for all future experiments. Among these two, nnU-net proved to perform slightly better.

The nnU-net is an extension of the U-Net in that it automates the intricate process of configuring the network parameters. This automation includes preprocessing, architecture selection, hyperparameter tuning, and postprocessing. The authors have introduced a systematic methodology that leverages heuristic principles and empirical data to adjust hyperparameters dynamically. This approach eliminates the tedious and time-consuming manual process of parameter tuning, which often requires a deep understanding of the domain but does not necessarily yield optimal results. Therefore, the nnU-net framework can perform various medical image segmentation tasks without requiring manual configuration or domain-specific knowledge. It offers two U-Net variants: the 2D U-Net, tailored for processing individual slices within a 3D volume, and the 3D U-Net, designed for processing entire 3D volumes. Both architectures include residual blocks, batch normalization and deep supervision techniques. In addition, nnU-net introduces a self-configuring data preprocessing pipeline that dynamically adjusts image properties such as orientation, spacing, cropping, and intensity normalization based on the unique characteristics of the dataset and the chosen network architecture. The model also employs an autonomous training strategy that optimizes critical parameters such as learning rate, batch size, number of training epochs, and data augmentation parameters by using a Bayesian optimization framework. nnU-net has achieved state-of-the-art results in several medical image segmentation benchmarks, such as the Medical Segmentation Decathlon [72], the KiTS19 Challenge [73], and the BraTS Challenge [74]. Due to its performance and user-friendly interface, this tool has since been widely adopted in the field of biomedical segmentation. Notably, it has been employed in various studies, including the segmentation of lung cancer in combination with GANs to generate training datasets [75], the segmentation of uteri in MRI data [76], as well as making a model for segmenting 104 different organs in CT data [77].

Swin UNETR builds on the U-Net architecture and incorporates the Swin transformer's attention mechanism, allowing the model to capture global contextual information across the entire image/volume. This attention mechanism enhances the model's ability to understand relationships between distant regions, unlike U-Net, which relies primarily on convolutional layers and might miss broader contextual connections. The Swin transformer that is used as the encoder in Swin UNETR creates a hierarchical feature representation, meaning that it captures features at varying resolutions and scales, which improves the model's ability to identify fine and course structures. The decoder used is U-Net-like, which knowingly excels in localizing segmentation features. The decoder merges multi-resolution features from the encoder using skip connections, providing high-resolution output while benefiting from the hierarchical understanding of the encoder. At it's publication, Swin UNETR marginally surpassed nnU-Net in the BraTS 2021 challenge (0.913 vs 0.908 Dice score on average). It has since been used for other segmentation tasks, such as tumor and lymph node segmentation [78].

To address the variable ultrasound image quality as already discussed in Section 1.1, we employed several strategies common in deep learning analysis. We used extensive data augmentation techniques during the training of our model, including random intensity variations and noise addition. These augmentations help the model generalize better to different image qualities. Additionally, the preprocessing pipeline of nnU-Net includes normalization steps to standardize the intensity distributions across different scans, mitigating the effects of varying image quality. We performed cross-validation to ensure consistent model performance across different data subsets, accounting for potential image quality variations. Our dataset also included volumes from multiple centers with different ultrasound machines and settings, exposing the model to a range of image qualities during training and evaluation.

We train both segmentation models to predict the location and segmentation of the uterine cavity in 3D ultrasound volumes. The models are trained to output a binary volume of the same size as the input, where a value of 1 depicts the uterine cavity, and 0 depicts everything else. To evaluate the performance, we use DSC, which measures the similarity between predicted and annotated segmentations, with higher values indicating better performance. We provide more details on the training procedures in Section 3.1.

2.3. Alignment

To enable a comparison between different 3D uterine shapes, we need to align the shapes to a common reference. For the alignment, we first tested traditional point set alignment methods such as ICP [79], normal distributions transform [80], and an algorithm based on phase correlation [81]. The results were unsatisfactory, mainly because uterine shapes vary significantly, while the methods were designed to align identical or similar shapes. We also decided not to use deep learning-based alignment approaches because, for unsupervised methods, such as the one presented by Zhou et al. [61], the dataset is not big enough, while for supervised deep learning approaches, such as the method presented by Shimada et al. [82], we lack the manual alignments.

We therefore propose a new alignment method based on geometric principles. The proposed method consists of several steps: selection of the target uterus 2.3.1, initial alignment of the uteri 2.3.2, curvature reduction 2.3.3, alignment of the uteri according to their principal axes 2.3.4, and fine-tuning of the alignment 2.3.5.

2.3.1. Choosing the target uterus

In our approach, each uterus is aligned to a manually selected target uterus (all-to-one alignment) to reduce the computational burden of all-to-all alignments and to allow easy updating when we receive new data. A qualitatively suitable specimen from the UterUS dataset was selected as the target, characterized by symmetry and a distinct cervix so that the rotations aren't biased in either direction. In addition, it has only a slight curvature so that the curvature reduction process doesn't significantly change its shape.

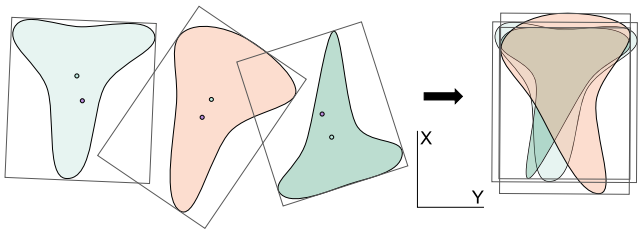


Fig. 6. A scheme of the initial step of the alignment process with the OBB shown in 2D for a clearer depiction. The purple dot shows the center of the bounding box (O) and the yellow dot shows the center of mass of the uterus (M).

2.3.2. Initial alignment

In the first alignment step, we roughly align the uteri by positioning and orienting their shapes in the same general direction. This means that the uterine bodies are in approximately the same position and the cervae point in approximately the same direction. This alignment is then gradually refined in the subsequent steps.

We calculate an approximate initial alignment of two uteri by aligning their bounding boxes. First, we look for the minimum oriented bounding box (OBB) with the smallest volume that completely encloses the uterus. We search for an oriented bounding box instead of an axis-aligned bounding box to ensure that the shape of the bounding box truly represents the shape of the uterus. The algorithm for finding OBB is based on convex hulls and principal component analysis (PCA) [83]. After calculating the bounding box of a selected uterus, it is aligned with the target by rotating it so that the axes of its box and the target box are parallel and by moving it so that the centers of the boxes coincide.

Due to the anatomy of the uterus and the imaging method used, the rotation of the uterus around the coronal (z) axis may be different for different volumes. Therefore, when aligning the bounding boxes, we must also check whether the aligned uterus should be rotated. Since the body of the uterus generally has more mass than the cervix, we compare the position of the center of mass of the uterus $M(M_x, M_y, M_z)$ with the center of the bounding box $O(O_x, O_y, O_z)$. If $M_x - O_x < 0$, we rotate the uterus by 180 degrees around the z axis (as is schematically shown with the rightmost green uterus shape in Fig. 6); otherwise, we leave it as it is.

2.3.3. Reducing the curvature

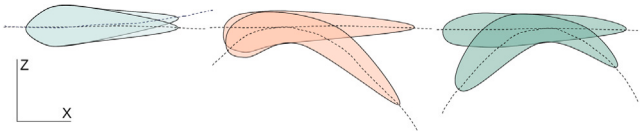


Fig. 7. Three examples of the curvature reduction step in 2D. The initial curvature surface and the flattened curvature surface are shown with dotted lines.

The uterine shapes within the body are curved to varying degrees, which makes precise alignment difficult. Moreover, this curvature has no anatomical significance. Gynecologists examining uterine shapes focus on the shape of the uterine cavity rather than its position and curvature in the body. Therefore, a reduction in uterine curvature is desirable, resulting in more accurate alignment.

We first estimate a curved surface, which we call the *curvature surface*, that best fits the uterus. This surface is then used to "flatten" the uterus and reduce its curvature. The schematic representation of the process in 2D can be seen in Fig. 7.

To find the curvature surface, we model the uterus shape as a point cloud and look for a curvature surface (plane) that best fits the points. We model the curvature surface with a quadratic function that can capture the overall curvature without overfitting the small bends. The curvature surface equation is given by:

$$f(x, y, \beta) \Rightarrow f(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F.$$

We estimate the parameters of the curvature surface $\beta = \{A, B, C, D, E, F\}$ by minimizing the sum of the squares of the distances between the point cloud coordinates P and the curvature surface:

$$\underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{m} \operatorname{dist}(P_i, f(x, y, \beta))^2$$

where m is the number of points in the point cloud and dist is the shortest distance between a point and the curvature surface. We use the Levenberg–Marquardt algorithm [84], an iterative algorithm that combines the Gauss–Newton method and the gradient descent method, to solve the nonlinear system of equations and estimate the curvature surface parameters β.

The obtained curvature surface is then used to reduce the curvature of the uterus in two steps:

1. *displacement calculation*, where we calculate the displacement of each point relative to the curvature surface and
2. *point cloud transformation*, where we move the points in the point cloud based on the calculated displacements.

For each point $P = (p_x, p_y, p_z)$ in the uterus point cloud, we find the closest point $S = (s_x, s_y, s_z)$ on the curvature surface by minimizing:

$$d(P, S) = (s_x - p_x)^2 + (s_y - p_y)^2 + (s_z - p_z)^2$$

where $s_z = f(s_x, s_y, \beta)$. We use the first and second derivatives of the function to identify critical points and estimate the closest point S.

In the second step, we create a new point cloud P' by reducing the curvature along the z axis for all points P in the point cloud as follows:

$$P'(x, y, z) = (p_x, p_y, p_z - s_z) = (p_x, p_y, p_z - f(s_x, s_y, \beta)).$$

2.3.4. Finding the principal axis

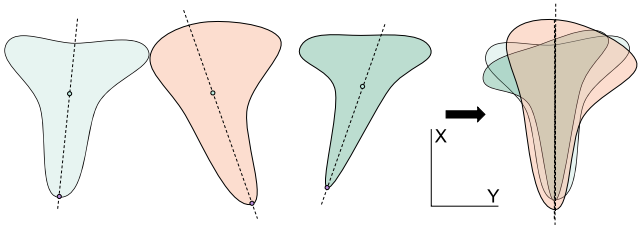


Fig. 8. A schematic of the alignment step using the principal axis in 2D.

When the uteri are flattened and thus better aligned along the z axis, they must also be aligned along the other two axes. This is accomplished by determining the principal axis of each uterus and aligning it with the corresponding axis of the target uterus. We define the principal axis as the axis connecting the midpoint of the uterine fundus (the upper part of the uterus, as shown in Fig. 1) to the midpoint of the lower part of the cervix. Since the cervix is usually more symmetrical than the body of the uterus, we approximate the principal axis by connecting the midpoint of the end of the cervix to the center of mass of the uterus (see Fig. 8).

To find the midpoint of the end of the cervix, we first project the point cloud of the uterus onto the XZ plane and fit a fourth degree polynomial curve to the 2D points to capture the curvature of the cervix. We then determine the last point on this curve that still overlaps with the uterus, which gives us the midpoint of the cervix in the XZ plane, denoted as $K_1 = (x_1, z_1)$. Next, we project the uterus onto the XY plane and find the center point $K_2 = (x_2, y_2)$ using a similar method. We combine these two points to get the midpoint of the cervix in 3D as $K = (x_2, y_2, z_1)$.

We use the resulting axis from K to the center of mass of the cervix to calculate the rotation angle that aligns the principal axis of the cervix with the principal axis of the target. We then transform all points in the point cloud of the uterus accordingly.

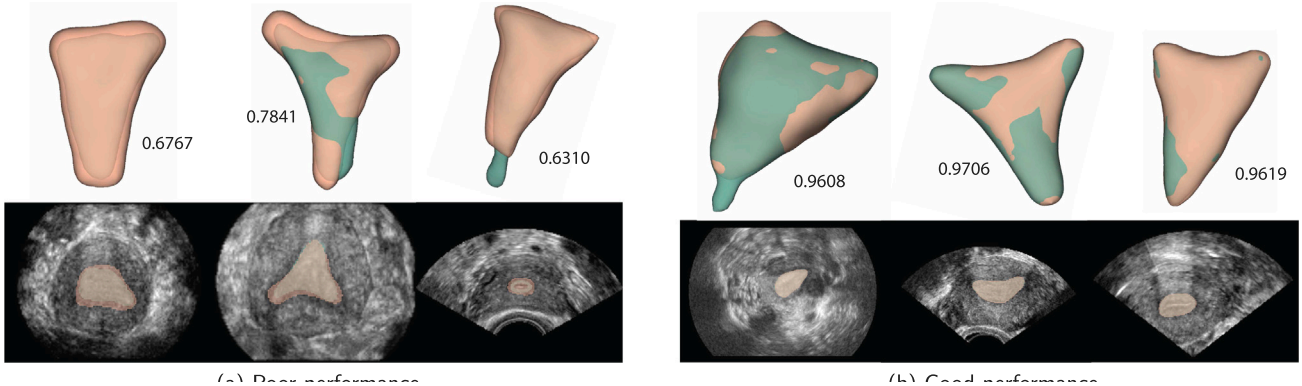


Fig. 9. Volumes with the worst and best evaluation scores with DSC values. The green color represents the prediction of our model, and the orange is the ground truth from our dataset.

2.3.5. Fine-tuning of rotation and translation

To improve the accuracy of the alignment, we use simulated annealing [85] to iteratively refine the alignment. Starting from a pair of approximately aligned uteri (from the previous step), we optimize six parameters representing the translation (t_x, t_y, t_z) and rotation (r_x, r_y, r_z) of the aligned uterus relative to the coordinate system of the target uterus

$$s = \{t_x, t_y, t_z, r_x, r_y, r_z\}.$$

We define the energy of the annealed system E as the *non-overlapping ratio (NOR)*:

$$E(\text{sample}, \text{target}) = \frac{|\text{combined}|}{|\text{sample}| + |\text{target}|}$$

where after sampling the point clouds into a uniform voxel grid, $|combined|$ represents the number of overlapping points and |sample| and |target| the number of points in the sampled and target uterus grids, respectively. A value of $NOR = 0.5$ represents a complete overlap, while $NOR = 1$ represents two non-overlapping point clouds. Therefore, our goal is to find a set of translation and rotation parameters that minimizes the energy function and thus increases the overlap of the two uteri.

Simulated annealing works iteratively, where in each iteration the state is changed by randomly selecting a parameter i and changing its value as follows:

$$s_i' = s_i + \xi * t_N$$

where ξ represents a random variable chosen uniformly from the interval $[-0.05, 0.05]$ for translation and $[-0.35, 0.35]$ for rotation. t_N represents the *neighbor temperature*, which is initially set to 100 and decreases by 1 every 100 iterations, fine-tuning the randomness of the parameters. At each step, the acceptance probability of a new state is defined as follows

$$P(\Delta, T) = \begin{cases} e^{-\frac{\Delta}{T}}, & \Delta \geq 0 \\ 1, & \Delta < 0 \end{cases}$$

where $\Delta = E' - E$ is the energy difference between the new state and the previous state and T is the current temperature of the system. T changes over the course of the algorithm according to the annealing schedule. In our process, we start at 0.01 and decrease it by $\alpha = 10^{-6}$ in each iteration until it reaches the value 0, at which point the algorithm terminates.

3. Results and discussion

3.1. Segmentation

Both networks were trained and evaluated on 141 annotated volumes from four different hospitals, all obtained with ultrasound machines from the same manufacturer (GE). The dataset was randomly divided into a training set (116 volumes) and a test set (28 volumes).

Table 1
Evaluation of the 3D nnU-Net and Swin UNETR segmentation models using the DSC metric on validation and test set.

	nnU-Net	Swin UNETR	
	3D	*pretrained*	*scratch*
Validation Set	0.909	0.860	0.851
Test Set	0.899	0.833	0.816

The training of nnU-Net was performed with five-fold cross-validation, where each model was trained five times, using 80% of the volumes of the training set for training and the remaining 20% for validation. Stochastic gradient descent with Nesterov momentum [86] with $\mu = 0.99$ was used for optimization, with an initial learning rate of 0.01. The learning rate decreased over epochs according to the *poly* policy [87], with the power factor set to 0.9. We ran two nnU-Net models: one operating only on 2D images and a full 3D model. The model trained on 2D images showed significantly worse performance (mean DSC of 0.8 compared to 0.9 with 3D on the validation set), highlighting the advantage of volumetric representation over 2D image processing, which is why we only evaluated the 3D model with other datasets. Besides 2D and 3D models, nnU-Net offers a 3D cascade model where first a 3D U-Net is trained on low-resolution images, and then a second high-resolution 3D U-Net refines the predictions of the former. Because the size of our volumes was relatively small, we have not used this as the low and high resolution models would be basically the same.

With Swin UNETR, we compared two models - one that we trained from scratch and a model pretrained on CT data [88]. We used the learning rate of 0.001 and employed the Adam optimizer [89] for stochastic gradient descent with the default values of $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 1e-8$.

In both cases, we used a compound loss function — a combination of cross-entropy loss and Dice loss [90] which has been shown to be more robust and work well on imbalanced data [91]. For the evaluation during and after training, the Dice coefficient (DSC) was used as the evaluation measure. The training was conducted using NVIDIA A100 GPU on the HPC system Vega [92].

The segmentation results in Table 1 show that the nnU-Net 3D model performs slightly better than both Swin UNETR models. While this seems somewhat inconsistent with the performance reported by Swin UNETR in comparison with nnU-Net in their original paper, the transformer-based models are known for requiring a lot of data in order to generalize well and not overfit. This is why it performs worse than nnU-Net in our case where the available data is not as extensive. The nnU-Net model performs well on average, with no large differences between the validation and test datasets, indicating that the model did not overfit on the training data. Additionally, we evaluated the nnU-Net model on 10 newly annotated volumes from a hospital not

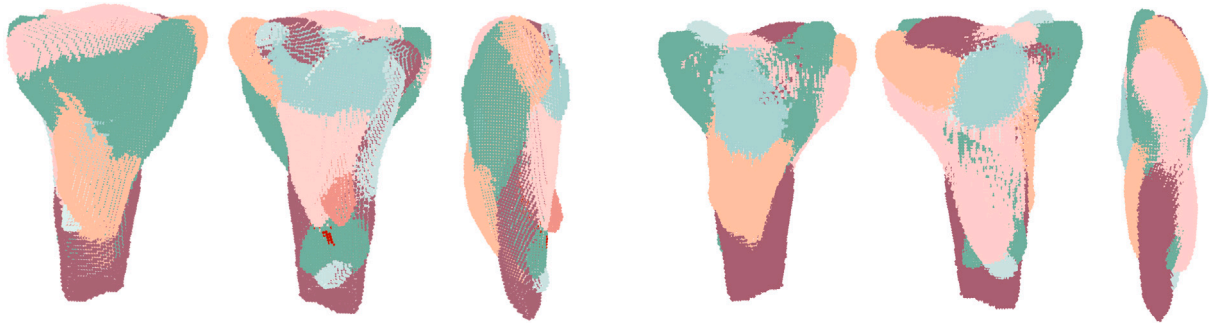


Fig. 10. Visual results of the alignment on 10 uterine shapes.

Table 2
Differences in translations and rotations of automatically and manually aligned uteri. Rotations are given in degrees and translations in voxels.

	our	ICP	CPD
translation X	3.31	7.34	4.31
translation Y	1.25	2.83	0.87
translation Z	0.90	0.77	0.86
rotation X	11.30	79.48	82.06
rotation Y	3.01	4.94	4.50
rotation Z	5.34	61.39	48.33

included in the training or testing datasets. The model maintained a high performance with a DSC of 0.863, indicating its robustness even with data from new operators. These additional volumes were acquired using the same GE ultrasound machine as the training data.

Looking at the three volumes with the lowest evaluation scores in Fig. 9(a), the model seems to label the boundaries more narrowly than in the manual annotations. The reason for this could be that, due to the poorer quality of the scans, the annotator preferred to mark the boundaries wide rather than too narrow when they were difficult to recognize in the cross-sections. In the third example, the model labeled the cervix while the annotator did not. We could even see this as an improvement, as localizing the cervix was often difficult during manual labeling because the image quality of the ultrasound decreases the further you get from the uterine cavity. In addition, there can be disagreement even among medical professionals about the exact position of the end of the cervix, which can lead to discrepancies in labeling. Therefore, the segmentations shown in the figure are not necessarily incorrect, even if the evaluation scores are low. Fig. 9(b) shows the volumes with the highest DSC values. It can be seen that the image quality of the scans is better, the volumes have less noise and the borders of the uterus are easier to distinguish from the background, which is why the manual annotations and the automatic segmentations largely match.

While the reported scores provide an indication of our approach's performance, they do not fully capture its true effectiveness and usability. We acknowledge the potential bias in our manual annotations, especially given that the annotations were performed by a single annotator. To address this, we plan to implement a more comprehensive annotation process involving multiple annotators. This will help to mitigate bias and provide a more robust evaluation of our method.

3.2. Alignment

The automatic alignment method was evaluated by qualitatively comparing a set of 10 automatically aligned uteri with the same set of manually aligned ones. The manual alignment was made with the assistance of a domain expert. Fig. 10 presents a visual comparison between the automatic and manual alignments from different perspectives. For quantitative comparison, we also evaluated two traditional shape alignment approaches: Iterative Closest Point (ICP) and Coherent Point Drift (CPD). Table 2 presents the average differences in translations and rotations calculated between each pair of automatically and manually aligned uteri for our approach, ICP, and CPD. The reported values are based on the same 10 uteri depicted in Fig. 10.

We can see that our approach produced good results both visually and numerically. The visualization in Fig. 10 shows that the result of automatic alignment is very similar to manual alignment. The main difference can be observed in the side view, where curvature reduction (which was not performed for manual alignment) improved alignment along the coronal axis. Reduction of the curvature increased the overlap and smoothness of the combined shape.

Similarly, the numerical results in Table 2 indicate that translation differences (measured in voxels) are minimal across all axes. Although looking at translations, the differences between the methods are not substantial, the advantage of our approach is evident in the rotation results, where the average degrees of difference are significantly smaller. The rotations along Y and Z axes are small as the initial bounding box alignment and subsequent refinements align those axes very well. A somewhat larger difference can be seen on the X axis rotation - the difference mostly arises when aligning smaller uteri inside a bigger one, as our current optimization metric in the final alignment step does not discern between the quality of alignment when one of the shapes is completely encapsulated in the other. Thus, final fine-tuning can, in some cases, worsen the result as it finds non-optimal solutions that are metric-wise equally good.

Manual alignment also allowed for more refined adjustments because we could simultaneously align all of the shapes (not just individual pairs), leading to a seemingly more connected surface across the combined uterus. Our pairwise approach can lead to suboptimal results as a sample might better align with the target uterus but worse with all of the others. However, modifying the algorithm to work on an all-to-all basis instead of all-to-one would significantly increase computational time. Moreover, it would necessitate repeating the entire alignment process when new data are introduced — with our current method, we only need to calculate the alignments of the new uteri to the target uterus and incorporate the results.

There is room for improvement of the proposed alignment method. When reducing the curvature, we only displace points on the z-axis, which can result in some deformations, especially when curvature is high. However, most of the samples in our dataset exhibit low or moderate levels of curvature, so the deformations don't have a large influence on the final result. Fig. 11 illustrates curve reduction in two uteri with low to moderate levels of curvature and the deformation caused by the flattening step. To improve our method further, we could also introduce improved optimization metrics in the final fine-tuning step, as only looking at the number of non-overlapping points is not

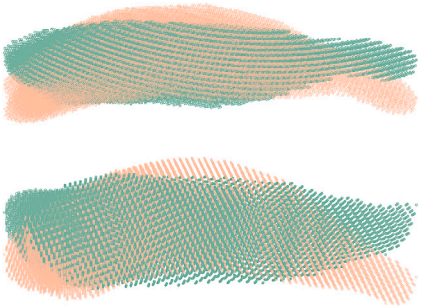


Fig. 11. Uteri point clouds and their flattened shapes after curvature reduction. Original point clouds are presented with orange and flattened shapes with green color.

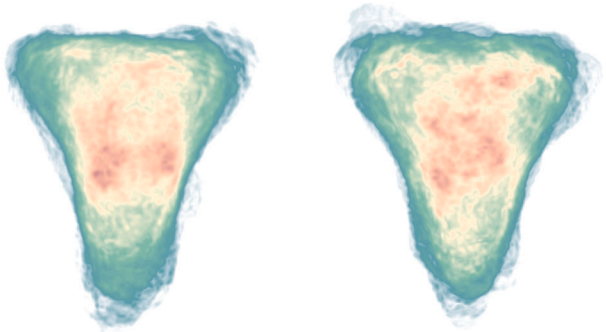


Fig. 12. Visualization of the overall shapes of the uterus for the general population (left) and the combined group of women with recurrent spontaneous miscarriages and unexplained infertility (right). The darker green color shows a lower density of boundary points, whereas the brighter yellow and orange present higher density.

optimal. We could consider a measure that incorporates more shape information, such as graph matching of the medial scaffold as presented by Chang et al. [93].

Alignment enables us to visualize the overall 3D uterine shape, as shown in Fig. 12. The visualization was done with Avizo3D [94] and shows the overall shape of the uterus for the general population (normal uteri) on the left and the overall shape of the remaining two groups (women with recurrent spontaneous miscarriages and women with unexplained infertility) on the right side.

The visualizations presented are not intended to draw definitive conclusions at this stage due to the limited data available. However, preliminary results indicate that the overall uterus shape resembles the letter *delta*, a characteristic typically observed in a healthy uterus. Our primary focus and the focus of the initial study wasn't on comparing different groups, but rather on establishing the shape of a normal uterus. With that in mind, the limited data available makes it difficult to draw distinctions between the two groups. Furthermore, the *others* group itself includes distinct subgroups with potentially opposing characteristics. Their inclusion, coupled with the limited data size, could easily mask individual variations, leading to an overall shape that is very similar to that of the general population. Fig. 13 visualizes the overall shape of automatically aligned uteri for both groups, highlighting areas of boundary density. Consistent with our expectations, the cervix exhibits the most pronounced variability, while the body-cervix region, particularly in the *general* population (red zones), demonstrates greater morphological consistency.

In future work, we plan to employ SSMs or some other approaches to derive more meaningful mean shapes than those obtained through simple averaging. This will enable us to capture and analyze shape variations within a population more effectively, providing deeper insights into the underlying anatomical differences and identifying and quantifying subtle shape differences that may be clinically significant.

Future efforts will also prioritize acquiring additional and less biased data, particularly for the *others* group, potentially subdividing it into its constituent subgroups. Only then will we be able to draw more precise conclusions about potential shape differences, classifications, and their clinical implications. This will enable the true potential of our alignment method in understanding and diagnosing uterine abnormalities, ultimately improving clinical outcomes.

4. Conclusion

In this study, we present an approach for the automatic segmentation and alignment of uteri in 3D ultrasound data, which can be a valuable tool for research and diagnostics in medicine. The main contributions of our work include a manually annotated dataset, a model for automatic uterus segmentation, and a novel approach for the automatic alignment of segmented 3D uteri.

To our knowledge, the presented segmentation model is the first of its kind for this type of data. The segmentations are accurate, with an average Dice coefficient of 0.863 on the test dataset. A qualitative investigation also confirmed the accuracy of the segmentations. The presented alignment algorithm can be used to observe the average uterine shape of the general population, which has already attracted great interest in the medical community. We have also made the annotated dataset of 3D ultrasound volumes (*UterUS*) available to the public, along with the trained segmentation model.

However, there is still plenty of room for improvement. We plan to enhance both the segmentation model and the final uterus reference model as more data is collected as part of the NURSE research initiative. We also aim to refine the manual annotations, as occasional errors still negatively impact the training of the segmentation model despite expert review. Involving multiple annotators and experts in the annotation process could improve the model's accuracy, although this work is time-consuming and costly. Additionally, as discussed in Section 3.2, a larger dataset will enable us to explore deep learning alignment methods that may prove more robust than the proposed approach.

There are numerous other promising avenues for future research in this field. These include gathering a more geographically diverse dataset, conducting longitudinal studies to observe how uterine shape changes over time, and incorporating other modalities such as 4D ultrasound. While these opportunities are full of potential, they are not our immediate priority at this time.

We hope that this work will lead to new insights in the field of gynecology. The initial feedback from participants in the NURSE multicenter study has been very positive. The preliminary results presented at the 32nd Annual Meeting of the ESGE (European Society for Gynecological Endoscopy) in Brussels show overlapping data on the shape of the uterine cavity in the general population. A new visualization and representation of the normal shape of the uterus would allow gynecologists to detect and surgically correct potential pathological abnormalities more quickly, potentially shortening the time to conception and reducing the rate of spontaneous miscarriages.

CRediT authorship contribution statement

Eva Boneš: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marco Gergolet:** Writing – review & editing, Validation, Formal analysis, Data curation, Conceptualization. **Ciril Bohak:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology. **Žiga Lesar:** Methodology, Conceptualization. **Matija Marolt:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

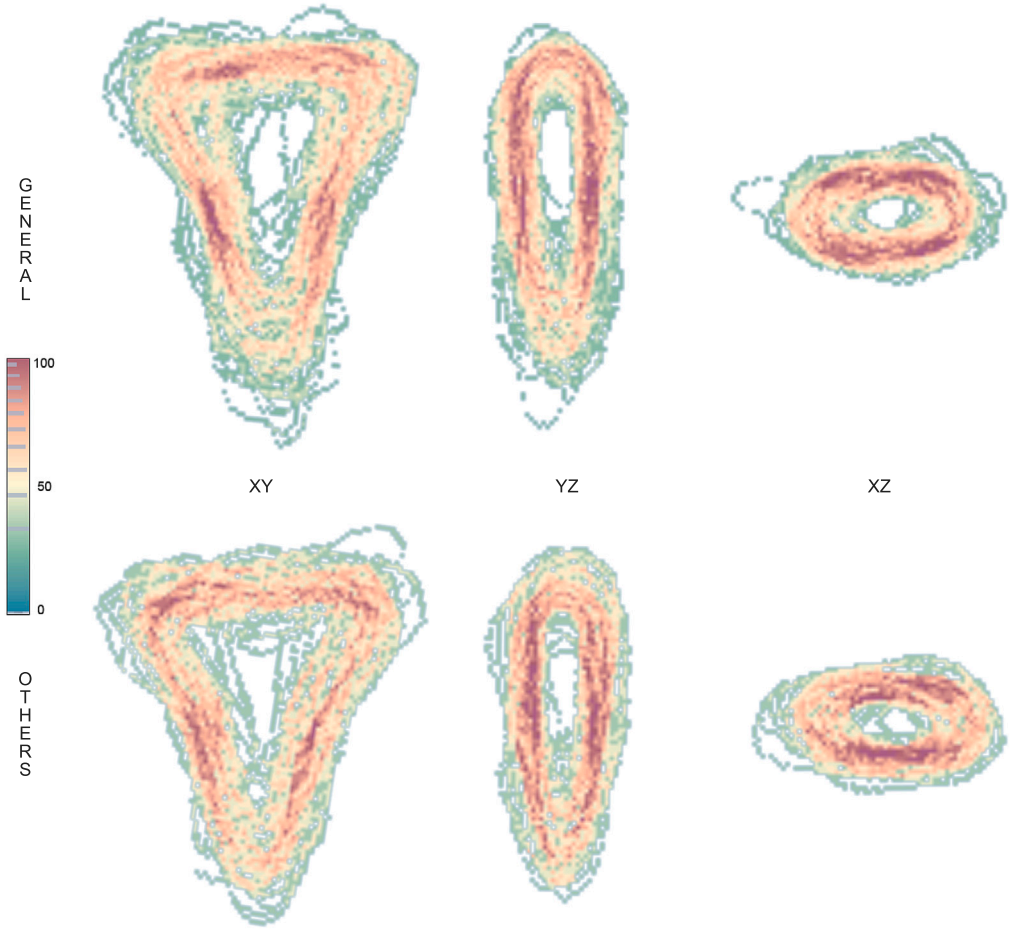


Fig. 13. Overall shape (overlayed boundaries of aligned uteri) for the *general* (on the top) and *others* (on the bottom) groups shown as an octo slice from XY, YZ, and XZ orientations (left, middle and right, respectively).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the following gynecologists for providing the data for our database: Marco Gergolet (Gynecology Clinic Gergolet), Maribel Acien (Miguel Hernández University of Elche), Alessandro Arena, Paolo Casadio (Alma Mater Studiorum - University of Bologna), Amerigo Vitagliano, Giuseppe Sorgente, Pierpaolo Nicoli (Gynecology Clinic at the University of Bari), Carmina Bermejo (Complutense University of Madrid), Attilio Di Spiezio Sardo, Brunella Zizolfi (University of Naples), Anna Biasioli, Lorenza Driul (Gynecology Clinic at the University of Udine), Anjeza Xholli, Umberto Scovazzi, Angelo Cagnacci (University of Genova), Stefano Bettocchi, Fabiana Fascilla (Gynecology Department at the General Hospital of Forli), Sotirios Saravelos (Imperial College London), Lubomir Mikulasek (General Hospital in Prague), Luca Savelli (Gynecology Department at the General Hospital of Forli), Francesca Buonomo, Clarice de Almeida Fiorillo, Martina Colombin (University of Trieste), Sara Brucker, Gertruda Jonaityte (University of Tübingen), Apostolos Athanasiadis, Anna Kougioumtsidou, Theodoris Theodoridis (Aristotle University of Thessaloniki), and Luis Alonso Pacheco (Clinica Victoria in Malaga). We also thank Michele Losappio, Marco De Pasquale, and Samsung's R&D team for all the help with converting their volumes for our usage.

References

[1] Douglas C. Wolf, Evaluation of the Size, Shape, and Consistency of the Liver, third ed., Butterworths, Boston, 1990.

[2] Paul J. Harrison, Nick Freemantle, John R. Geddes, Meta-analysis of brain weight in Schizophrenia, Schizophr. Res. 64 (1) (2003).

[3] H. Gray, Gray's Anatomy, thirtyninth ed., Churchill, London, 2005, pp. 997–1003.

[4] Peter Oppelt, Stefan P. Renner, Sara Brucker, Pamela L. Strissel, Reiner Strick, Patricia G. Oppelt, Hellmuth G. Doerr, Guenther E. Schott, Juergen Hucke, Diethelm Wallwiener, Matthias W. Beckmann, The VCUAM (Vagina cervix uterus adnex-associated malformation) classification: A new classification for genital malformations, Fertil. Steril. 84 (5) (2005).

[5] Grigoris F. Grimbizis, Stephan Gordts, Attilio Di Spiezio Sardo, Sara Brucker, Carlo De Angelis, Marco Gergolet, Tin Chiu Li, Vasilios Tanos, Hans Brölmann, Luca Gianaroli, Rudi Campo, The ESHRE/ESGE consensus on the classification of female genital tract congenital anomalies, Hum. Reprod. 28 (8) (2013).

[6] Marco Gergolet, Invited lecture - What is normal uterus: An update on the NURSE study, in: Proceedings of the ESGE 32nd Annual Congress, Brussels, Belgium, 2023.

[7] Fabian Isensee, Paul F. Jaeger, Simon A.A. Kohl, Jens Petersen, Klaus H. Maier-Hein, nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation, Nat. Methods 18 (2) (2021).

[8] James Nott, Anatomy of the uterus, female reproductive anatomy, geeky medics, 2020, Available at: https://geekymedics.com/anatomy-of-the-uterus/. (Accessed 17 June 2024).

[9] Harold Ellis, Anatomy of the uterus, Anaesth. Intensive Care 12 (3) (2011) 99–101.

[10] Chris T. Bolliger, Felix J.F. Herth, Paul H. Mayo, Teruomi Miyazawa, John F. Beamis, Clinical Chest Ultrasound: From the ICU to the Bronchoscopy Suite, in: Progress In Respiratory Research, S.Karger AG, 2009.

[11] Fikri M. Abu-Zidan, Ashraf F. Hefny, Peter Corr, Clinical ultrasound physics, J. Emerg. Trauma Shock 4 (4) (2011).

[12] Hyunyoung Ahn, Edgar Hernández-Andrade, Roberto Romero, Manasi Ptwardhan, Luis F. Goncalves, Alma Aurioles-Garibay, Maynor Garcia, Sonia S. Hassan, Lami Yeo, Mirror artifacts in obstetric ultrasound: Case presentation of a ghost twin during the second-trimester ultrasound scan, Fetal. Diagn. Ther. 34 (4) (2013).

[13] Barbara S. Hertzberg, William D. Middleton, Ultrasound the Requisites, Elsevier, 2016.

[14] T. Joel, R. Sivakumar, An extensive review on despeckling of medical ultrasound images using various transformation techniques, Appl. Acoust. 138 (2018) 18–27.

[15] Boris Tutschek, Three-dimensional ultrasound: Techniques and clinical applications, in: Obstetric Imaging: Fetal Diagnosis and Care, second ed., Elsevier, 2017, pp. 713–721.

[16] S. Lori Bridal, Jean Michel Correas, Amena Saïed, Pascal Laugier, Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging, Proc. IEEE 91 (10) (2003) 1543–1561.

[17] Mainak Biswas, Venkatanareshbabu Kuppili, Damodar Reddy Edla, Harman S. Suri, Luca Saba, Rui Tato Marinhoe, J. Miguel Sanches, Jasjit S. Suri, Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm, Comput. Methods Programs Biomed. 155 (2018).

[18] Zhantao Cao, Lixin Duan, Guowu Yang, Ting Yue, Qin Chen, An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures, BMC Med. Imaging 19 (1) (2019).

[19] Jianning Chi, Ekta Walia, Paul Babyn, Jimmy Wang, Gary Groot, Mark Eramian, Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network, J. Digit. Imaging 30 (4) (2017).

[20] Rosa María Menchón-Lara, José Luis Sancho-Gómez, Andrés Bueno-Crespo, Early-stage atherosclerosis detection using deep learning over carotid ultrasound images, Appl. Soft Comput. 49 (2016).

[21] Bin Pu, Kenli Li, Shengli Li, Ningbo Zhu, Automatic fetal ultrasound standard plane recognition based on deep learning and IIoT, IEEE Trans. Ind. Inf. 17 (11) (2021) 7771–7780.

[22] Xin Yang, Lequan Yu, Shengli Li, Huaxuan Wen, Dandan Luo, Cheng Bian, Jing Qin, Dong Ni, Pheng-Ann Heng, Towards automated semantic segmentation in prenatal volumetric ultrasound, IEEE Trans. Med. Imaging 38 (1) (2019) 180–193.

[23] Ryan Cunningham, María Sánchez, Ian Loram, Ultrasound segmentation of cervical muscle during head motion: A dataset and a benchmark using deconvolutional neural networks, 2019, http://dx.doi.org/10.31224/osf.io/fsa3c, PrePrint.

[24] Ran Zhou, Aaron Fenster, Yujiao Xia, J. David Spence, Mingyue Ding, Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images, Med. Phys. 46 (7) (2019) 3180–3193.

[25] Xin Yang, Haoran Dou, Ran Li, Xu Wang, Cheng Bian, Shengli Li, Dong Ni, Pheng-Ann Heng, Generalizing deep models for ultrasound image segmentation, in: Proceedings of the Medical Image Computing and Computer Assisted Intervention, MICCAI, Springer International Publishing, 2018, pp. 497–505.

[26] Evan Shelhamer, Jonathan Long, Trevor Darrell, Fully convolutional networks for semantic segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (4) (2017).

[27] Yizhe Zhang, Michael T.C. Ying, Lin Yang, Anil T. Ahuja, Danny Z. Chen, Coarse-to-fine stacked fully convolutional nets for lymph node segmentation in ultrasound images, in: Proceedings of the International Conference on Bioinformatics and Biomedicine, BIBM, 2017, pp. 443–448.

[28] Lingyun Wu, Yang Xin, Shengli Li, Tianfu Wang, Pheng Ann Heng, Dong Ni, Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation, in: Proceedings of the International Symposium on Biomedical Imaging, ISBI, IEEE, 2017, pp. 663–666.

[29] Deepak Mishra, Santanu Chaudhury, Mukul Sarkar, Arvinder Singh Soin, Ultrasound image segmentation: A deeply supervised network with attention to boundaries, IEEE Trans. Biomed. Eng. 66 (6) (2019).

[30] Chen-Yu Lee, Saining Xie, Patrick Gallagher, Zhengyou Zhang, Zhuowen Tu, Deeply-supervised nets, in: Proceedings of the International Conference on Artificial Intelligence and Statistics, Vol. 38, AISTATS, PMLR, 2015, pp. 562–570.

[31] Olaf Ronneberger, Philipp Fischer, Thomas Brox, U-Net: Convolutional networks for biomedical image segmentation, in: Proceedings of Medical Image Computing and Computer Assisted Intervention, Vol. 9351, MICCAI, Springer International Publishing, 2015, pp. 234–241.

[32] Jeya Maria Jose Valanarasu, Vishwanath A. Sindagi, Ilker Hacihaliloglu, Vishal M. Patel, KiU-Net: Towards accurate segmentation of biomedical images using over-complete representations, in: Proceedings of Medical Image Computing and Computer Assisted Intervention, MICCAI, 2020, pp. 363–373.

[33] Pankaj K. Jain, Neeraj Sharma, Argiris A. Giannopoulos, Luca Saba, Andrew Nicolaides, Jasjit S. Suri, Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound, Comput. Biol. Med. 136 (2021).

[34] Viksit Kumar, Jeremy M. Webb, Adriana Gregory, Max Denis, Duane D. Meixner, Mahdi Bayat, Dana H. Whaley, Mostafa Fatemi, Azra Alizad, Automated and real-time segmentation of suspicious breast masses using convolutional neural network, Plos One 13 (5) (2018).

[35] Lu Xu, Shengbo Gao, Lijuan Shi, Boxuan Wei, Xiaowei Liu, Jicong Zhang, Yihua He, Exploiting vector attention and context prior for ultrasound image segmentation, Neurocomputing 454 (2021).

[36] Ozan Oktay, Jo Schlemper, Loïc Le Folgoc, Matthew C. H. Lee, Mattias P. Heinrich, Kazunari Misawa, Kensaku Mori, Steven G. McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, Daniel Rueckert, Attention U-net: Learning where to look for the pancreas, in: Proceedings of Medical Imaging with Deep Learning, MIDL, 2018.

[37] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, Xiaoou Tang, Residual attention network for image classification, in: Proceedings of Computer Vision and Pattern Recognition, CVPR, IEEE Computer Society, 2017, pp. 6450–6458.

[38] Junlong Cheng, Shengwei Tian, Long Yu, Chengrui Gao, Xiaojing Kang, Xiang Ma, Weidong Wu, Shijia Liu, Hongchun Lu, ResGANet: Residual group attention network for medical image classification and segmentation, Med. Image Anal. 76 (2022) 102313.

[39] Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R. Roth, Daguang Xu, Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images, in: Proceedings of Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, BrainLes, Springer International Publishing, 2022, pp. 272–284.

[40] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the International Conference on Computer Vision, ICCV, IEEE Computer Society, 2021, pp. 9992–10002.

[41] Bahareh Behboodi, Hassan Rivaz, Susan Lalondrelle, Emma Harris, Automatic 3D ultrasound segmentation of uterus using deep learning, in: Proceedings of the International Ultrasonics Symposium, IUS, IEEE, 2021.

[42] Maysam Shahedi, James Dormer, Quyen Do, Yin Xi, Matthew Lewis, Christina Herrera, Catherine Spong, Ananth Madhuranthakam, Diane Twickler, Baowei Fei, Automatic segmentation of uterine cavity and placenta on MR images using deep learning, in: Proceedings of Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging, Vol. 12036, International Society for Optics and Photonics, 2022.

[43] Bin Kong, Xin Wang, Yi Lu, Hao Yu Yang, Kunlin Cao, Qi Song, Youbing Yin, Ultrasonography uterus and fetus segmentation with constrained spatial-temporal memory FCN, in: Proceedings of Medical Image Understanding and Analysis, Vol. 13413 LNCS, MIUA, Springer International Publishing, 2022, pp. 253–261.

[44] Yasuhisa Kurata, Mizuho Nishio, Aki Kido, Koji Fujimoto, Masahiro Yakami, Hiroyoshi Isoda, Kaori Togashi, Automatic segmentation of the uterus on MRI using a convolutional neural network, Comput. Biol. Med. 114 (2019) 103438.

[45] Lee R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945).

[46] T. Sørensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons, Biologiske Skrifter 5 (1948).

[47] Qing Zhou, Yunhao Bai, Fang Chen, He Zhang, Limei Chen, Guofu Zhang, Yan Wang, Long Sui, 3D reconstruction of the uterus and automatic segmentation of the uterine cavity on 3D magnetic resonance imaging: A preliminary study, Heliyon 10 (1) (2024) e23558.

[48] Lingxi Xie, Qihang Yu, Yuyin Zhou, Yan Wang, Elliot K. Fishman, Alan L. Yuille, Recurrent saliency transformation network for tiny target segmentation in abdominal CT scans, IEEE Trans. Med. Imaging 39 (2) (2020) 514–525.

[49] Timo Mäkelä, Patrick Clarysse, Outi Sipilä, Nicoleta Pauna, Quoc Cuong Pham, Toivo Katila, Isabelle E. Magnin, A review of cardiac image registration methods, IEEE Trans. Med. Imaging 21 (9) (2002) 1011–1021.

[50] Jiaolong Yang, Hongdong Li, Dylan Campbell, Yunde Jia, Go-ICP: A globally optimal solution to 3D ICP point-set registration, IEEE Trans. Pattern Anal. Mach. Intell. 38 (11) (2016).

[51] Andriy Myronenko, Xubo Song, Miguel Carreira-Perpiñán, Non-rigid point set registration: Coherent point drift, in: Proceedings of the Advances in Neural Information Processing Systems, Vol. 19, NeurIPS, Curran Associates, Inc., 2006, pp. 1009–1016.

[52] Mohamed Chaouch, Anne Verroust-Blondet, Alignment of 3D models, Graph. Models 71 (2) (2009) 63–76.

[53] Robert Cupec, Ivan Vidović, Damir Filko, Petra Djurović, Object recognition based on convex hull alignment, Pattern Recognit. 102 (2020).

[54] David G. Lowe, Object recognition from local scale-invariant features, in: Proceedings of the International Conference on Computer Vision, Vol. 2, ICCV, IEEE, 1999, pp. 1150–1157.

[55] Herbert Bay, Tinne Tuytelaars, Luc Van Gool, SURF: Speeded up robust features, in: Proceedings of the European Conference on Computer Vision, ECCV, Springer Berlin Heidelberg, 2006, pp. 404–417.

[56] Martin A. Fischler, Robert C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981).

[57] Tobias Heimann, Hans-Peter Meinzer, Statistical shape models for 3D medical image segmentation: A review, Med. Image Anal. 13 (4) (2009) 543–563.

[58] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models-their training and application, Comput. Vis. Image Underst. 61 (1) (1995) 38–59.

[59] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, IEEE Trans. Pattern Anal. Mach. Intell. 23 (6) (2001) 681–685.

[60] Max Jaderberg, Karen Simonyan, Andrew Zisserman, Koray Kavukcuoglu, Spatial transformer networks, in: Proceedings of the Advances in Neural Information Processing Systems, Vol. 28, NeurIPS, Curran Associates, Inc., 2015.

[61] Keyang Zhou, Bharat Lal Bhatnagar, Bernt Schiele, Gerard Pons-Moll, Adjoint rigid transform network: Task-conditioned alignment of 3D shapes, in: Proceedings of the International Conference on 3D Vision, 3DV, IEEE Computer Society, 2022, pp. 1–11.

[62] Sotirios H. Saravelos, Tin Chiu Li, Intra-cycle variation of the uterine cavity indentation assessed with three-dimensional ultrasound in natural and stimulated cycles, Reprod. Biomed. Online 32 (5) (2016).

[63] Pádraig Looney, Gordon Stevenson, Sally Collins, 3D ultrasound file reading and coordinate transformations, J. Open Source Softw. 4 (33) (2019) 1063.

[64] Ron Kikinis, Steve D. Pieper, Kirby G. Vosburgh, 3D slicer: A platform for subject-specific image analysis, visualization, and clinical support, in: Intraoperative Imaging and Image-Guided Therapy, Springer New York, 2014, pp. 277–289.

[65] Andras Lasso, Christian Herz, Hannah Nam, Alana Cianciulli, Steve Pieper, Simon Drouin, Csaba Pinter, Samuelle St-Onge, Chad Vigil, Stephen Ching, Kyle Sunderland, Gabor Fichtinger, Ron Kikinis, Matthew A. Jolley, SlicerHeart: An open-source computing platform for cardiac image analysis and modeling, Front. Cardiovasc. Med. 9 (2022).

[66] Robert W Cox, John Ashburner, Hester Breman, Kate Fissell, Christian Haselgrove, Colin J Holmes, Jack L Lancaster, David E Rex, Stephen M Smith, Jeffrey B Woodward, Stephen C Strother, A (sort of) new image data format standard: Nifti-1, in: 10th Annual Meeting of the Organization for Human Brain Mapping, Vol. 22, OHBM, 2004.

[67] Megan Maar, Juhyun Lee, Anthony Tardi, Yuan Yi Zheng, Candance Wong, Jing Gao, Inter-transducer variability of ultrasound image quality in obese adults: Qualitative and quantitative comparisons, Clin. Imaging 92 (2022).

[68] Ted Lucidi, Noise artifacts in diagnostic ultrasound - Innovatus imaging, 2021, Available at: https://www.innovatusimaging.com/whats-that-noise-part-1/. (Accessed 17 June 2024).

[69] Manca Žerovnik Mekuč, Ciril Bohak, Samo Hudoklin, Byeong Hak Kim, Rok Romih, Min Young Kim, Matija Marolt, Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data, Comput. Biol. Med. 119 (2020).

[70] Sihong Chen, Kai Ma, Yefeng Zheng, Med3D: Transfer learning for 3D medical image analysis, 2019, http://dx.doi.org/10.48550/arXiv.1904.00625, PrePrint.

[71] Krishna Chaitanya, Ertunc Erdil, Neerav Karani, Ender Konukoglu, Contrastive learning of global and local features for medical image segmentation with limited annotations, in: Proceedings of the Advances in Neural Information Processing Systems, Vol. 33, NeurIPS, Curran Associates, Inc., 2020, pp. 12546–12558.

[72] Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A. Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M. Summers, Bram van Ginneken, Michel Bilello, Patrick Bilic, Patrick F. Christ, Richard K.G. Do, Marc J. Gollub, Stephan H. Heckers, Henkjan Huisman, William R. Jarnagin, Maureen K. McHugo, Sandy Napel, Jennifer S.Golia Pernicka, Kawal Rhode, Catalina Tobon-Gomez, Eugene Vorontsov, James A. Meakin, Sebastien Ourselin, Manuel Wiesenfarth, Pablo Arbeláez, Byeonguk Bae, Sihong Chen, Laura Daza, Jianjiang Feng, Baochun He, Fabian Isensee, Yuanfeng Ji, Fucang Jia, Ildoo Kim, Klaus Maier-Hein, Dorit Merhof, Akshay Pai, Beomhee Park, Mathias Perslev, Ramin Rezaiifar, Oliver Rippel, Ignacio Sarasua, Wei Shen, Jaemin Son, Christian Wachinger, Liansheng Wang, Yan Wang, Yingda Xia, Daguang Xu, Zhanwei Xu, Yefeng Zheng, Amber L. Simpson, Lena Maier-Hein, M. Jorge Cardoso, The medical segmentation decathlon, Nat. Commun. 13 (1) (2022).

[73] Nicholas Heller, Fabian Isensee, Klaus H. Maier-Hein, Xiaoshuai Hou, Chunmei Xie, Fengyi Li, Yang Nan, Guangrui Mu, Zhiyong Lin, Miofei Han, Guang Yao, Yaozong Gao, Yao Zhang, Yixin Wang, Feng Hou, Jiawei Yang, Guangwei Xiong, Jiang Tian, Cheng Zhong, Jun Ma, Jack Rickman, Joshua Dean, Bethany Stai, Resha Tejpaul, Makinna Oestreich, Paul Blake, Heather Kaluzniak, Shaneabbas Raza, Joel Rosenberg, Keenan Moore, Edward Walczak, Zachary Rengel, Zach Edgerton, Ranveer Vasdev, Matthew Peterson, Sean McSweeney, Sarah Peterson, Arveen Kalapara, Niranjan Sathianathen, Nikolaos Papanikolopoulos, Christopher Weight, The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge, Med. Image Anal. 67 (2021).

[74] Huan Minh Luu, Sung Hong Park, Extending nn-unet for brain tumor segmentation, in: Proceedings of Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, BrainLes, Springer International Publishing, 2022, pp. 173–186.

[75] Mizuho Nishio, Koji Fujimoto, Hidetoshi Matsuo, Chisako Muramatsu, Ryo Sakamoto, Hiroshi Fujita, Lung cancer segmentation with transfer learning: Usefulness of a pretrained model constructed from an artificial dataset generated using a generative adversarial network, Front. Artif. Intell. 4 (2021).

[76] Daniel Hausmann, Aline Lerch, Sebastian Hitziger, Monika Farkas, Elisabeth Weiland, Andreas Lemke, Maximilian Grimm, Rahel A. Kubik-Huch, AI-supported autonomous uterus reconstructions: First application in MRI using 3D SPACE with iterative denoising, Acad. Radiol. 31 (4) (2024) 1400–1409.

[77] Jakob Wasserthal, Hanns-Christian Breit, Manfred T. Meyer, Maurice Pradella, Daniel Hinck, Alexander W. Sauter, Tobias Heye, Daniel T. Boll, Joshy Cyriac, Shan Yang, Michael Bach, Martin Segeroth, TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images, Radiol. Artif. Intell. 5 (5) (2023).

[78] Hung Chu, Luis Ricardo De la O Arévalo, Wei Tang, Baoqiang Ma, Yan Li, Alessia De Biase, Stefan Both, Johannes Albertus Langendijk, Peter van Ooijen, Nanna Maria Sijtsema, Lisanne V. van Dijk, Swin UNETR for tumor and lymph node segmentation using 3D PET/CT imaging: A transfer learning approach, in: Proceedings of Head and Neck Tumor Segmentation and Outcome Prediction, HECKTOR, Springer Nature Switzerland, 2023, pp. 114–120.

[79] K.S. Arun, T.S. Huang, S.D. Blostein, Least-squares fitting of two 3-D point sets, IEEE Trans. Pattern Anal. Mach. Intell. PAMI (5) (1987).

[80] Peter Biber, Wolfgang Straßer, The normal distributions transform: A new approach to laser scan matching, in: Proceedings of the International Conference on Intelligent Robots and Systems, Vol. 3, IROS, 2003, pp. 2743–2748.

[81] Martin Dimitrievski, David Van Hamme, Peter Veelaert, Wilfried Philips, Robust matching of occupancy maps for odometry in autonomous vehicles, in: Proceedings of the Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Vol. 3, VISIGRAPP, 2016, pp. 628–635.

[82] Soshi Shimada, Vladislav Golyanik, Edgar Tretschk, Didier Stricker, Christian Theobalt, DispVoxNets: Non-rigid point set alignment with supervised learning proxies, in: Proceedings of the International Conference on 3D Vision, 3DV, IEEE Computer Society, 2019, pp. 27–36.

[83] Stefan Gottschalk, Dinesh Manocha, Ming C. Lin, Collision Queries Using Oriented Bounding Boxes (Ph.D. thesis), The University of North Carolina at Chapel Hill, 2000.

[84] Kenneth Levenberg, A method for the solution of certain non-linear problems in least squares, Q. Appl. Math. 2 (2) (1944) 164–168.

[85] Scott Kirkpatrick, D. Jr., Mario Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.

[86] Yurii Nesterov, A method of solving a convex programming problem with convergence rate $O(1/k^2)$, Dokl. Akad. Nauk SSSR 269 (3) (1983) 543.

[87] Liang Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille, DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2018).

[88] Yucheng Tang, Dong Yang, Wenqi Li, Holger R Roth, Bennett Landman, Daguang Xu, Vishwesh Nath, Ali Hatamizadeh, Self-supervised pre-training of swin transformers for 3D medical image analysis, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR, 2022, pp. 20730–20740.

[89] Diederik P. Kingma, Jimmy Ba, Adam: A method for stochastic optimization, in: Proceedings of the International Conference on Learning Representations, ICLR, 2015.

[90] Michal Drozdzal, Eugene Vorontsov, Gabriel Chartrand, Samuel Kadoury, Chris Pal, The importance of skip connections in biomedical image segmentation, in: Proceedings of Deep Learning and Data Labeling for Medical Applications, Vol. 10008, DLMIA LABELS, Springer International Publishing, 2016, pp. 179–187.

[91] Jun Ma, Jianan Chen, Matthew Ng, Rui Huang, Yu Li, Chen Li, Xiaoping Yang, Anne L. Martel, Loss odyssey in medical image segmentation, Med. Image Anal. 71 (2021).

[92] HPC RIVR Consortium, HPC system vega, 2022, Available at: http://www.hpc-rivr.si. (Accessed 17 June 2024).

[93] Ming Ching Chang, Benjamin B. Kimia, Measuring 3D shape similarity by graph-based matching of the medial scaffolds, Comput. Vis. Image Underst. 115 (5) (2011).

[94] Thermo Fisher Scientific, Avizo software for 3D visualization and analysis, 2023, Version 2023.1.1, https://www.thermofisher.com/sa/en/home/electron-microscopy/products/software-em-3d-vis/avizo-software.html. (Accessed 6 November 2023).