

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, 1000, Ljubljana, Slovenia*

^b*University of Ljubljana, Faculty of Medicine, Vrazov trg 2, 1000, Ljubljana, Slovenia*

^c*King Abdullah University of Science and Technology, Visual Computing Center, 23955-6900, Thuwal, Saudi Arabia*

Abstract

We present a novel pipeline for automatic 3D segmentation, alignment, and visualization of uterine shapes from ultrasound volumes. This work aims to establish a standardized model of the normal uterus, with the potential for advancements in women's health research. To achieve this, we built a large, annotated 3D ultrasound image dataset, facilitating robust segmentation model development and future research. Our nnU-net-based 3D segmentation model achieves Dice coefficient of 0.90 and demonstrates adaptability across data sources from different hospitals and technicians. Furthermore, we introduce a novel automated alignment method for 3D uterine shapes, employing mathematical and geometric techniques. This approach enables visualization and further detailed expert analysis of uterine morphology, potentially contributing to earlier detection of uterine abnormalities, improved diagnosis, treatment, and education in women's health.

Keywords: uterus segmentation, volumetric ultrasound, 3D alignment

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 [65, 29, 23], there is no large-scale study of the uterus that examines its normal characteristics. Some uterine shape classifications, such as VCUAM by [56] and ESHRE/ESGE by [28], 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 defined the normal uterus as an organ having "either straight or curved interstitial line but with an internal indentation at the fundal midline not exceeding 50% of the uterine wall thickness", avoiding to use absolute numbers. Our research is part of a larger study called NURSE (Normal UteRine asSEssment) 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.

*Corresponding author

Email addresses: eva.bones@lgu.fri.uni-lj.si (Eva Boneš), marco.gergolet@gmail.com (Marco Gergolet), ciril.bohak@fri.uni-lj.si (Ciril Bohak), ziga.lesar@fri.uni-lj.si (Žiga Lesar), matija.marolt@fri.uni-lj.si (Matija Marolt)

Although professionals can make uterine measurements manually by defining a cross-section of the 3D volume and annotating them on the resulting 2D image, such measurements are inherently limited by the obtained 2D view and prone to human error and inter-operator variability in the “virtual organ computer-aided analysis (VO-CAL)” application. To alleviate these limitations, we propose a method that extracts and aligns 3D uterine shapes from 3D ultrasound volumes, enabling us to create a reference model of a normal 3D uterus and compare 3D uterine shapes from different patients. Such aligned 3D views can provide additional information that manual 2D measurements cannot. By automating the entire process, the proposed method reduces the need for human involvement, reduces human error, and enables the processing of large amounts of 3D volumes in a fraction of the time required for manual annotation and alignment. In addition, the resulting 3D shapes allow for easy visualization, making them a useful tool for educational purposes.

The main contributions of our paper are:

1. We provide a neural model for **automatic segmentation** of uterine cavities in 3D ultrasound data. The model is based on the nnU-net architecture [32] and gives good results on 3D ultrasound data acquired in various hospitals by different operators.
2. We present and make available 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

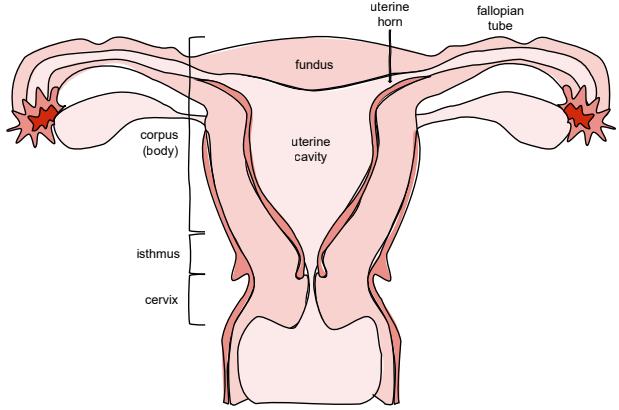


Figure 1: Rough anatomy of the uterus. Image inspired by [55].

uterine cavity; the isthmus – the part where the uterine cavity starts to narrow; and the uterine cervix. Figure 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 Figure 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 Figure 2. This curvature, among other complexities, poses a challenge for the comparison and automatic alignment of the uteri.

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 [9, 1, 2, 51]. Despite the noise, medical professionals are trained and skilled at telling apart artifacts from actual structures, but it is challenging for the automated approaches to do the same [35].

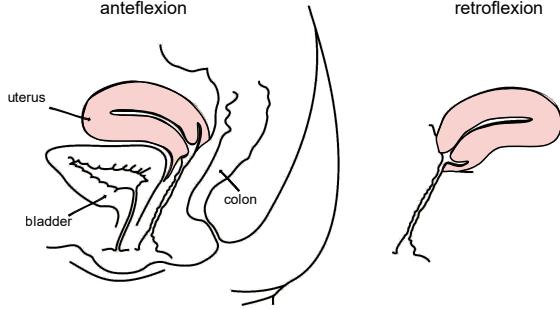


Figure 2: Position of the uterus in the body. Image adapted from [25].

3D ultrasound volumes are created by capturing a sequence of 2D images, either manually or automatically, using specialized 3D ultrasound probes that allow the direction adjustment of the ultrasound beam. The computer system records both the images and the spatial data required to construct the 3D volume. Accurate positioning is achieved by an electromagnetic position sensor or an electrical gyroscope integrated into the ultrasound probe. The quality of 3D ultrasound volumes is highly dependent on the correct placement of the ultrasound probe and the skill of the medical professional conducting the examination. Unlike many other imaging modalities, where the patient's position and the imaging direction are predetermined, ultrasound examination requires the operator to skillfully manipulate the probe to obtain optimal images and, thus, optimally constructed volumes. Ultrasound does not provide uniform resolution in all axes, so the imaging plane (acquisition plane) that is typically displayed on the ultrasound screen after volume acquisition offers the best visualization. Ensuring that the structure of interest, in our case, the uterus, is well represented in this initial plane is critical to obtaining accurate final results [63].

1.2. Related work

1.2.1. Ultrasound segmentation

As discussed in Section 1.1, ultrasound images present a significant challenge to automated analysis, from noise and artifacts resulting from the physical properties of ul-

trasound propagation to the fact that the quality of the images is highly dependent on the operator performing the ultrasound. Recently, with the technological advances in the development of ultrasound equipment, such as improved ultrasound probes, higher resolution, new digital systems, etc., the quality of ultrasound images has improved [10]. In addition, the development of deep learning has enabled more successful data analysis, making this field highly active.

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 [8], breast [11], thyroid [17], and arteries [50] to recurrent neural networks (RNN)[57, 71], encoder-decoder networks [19, 76], and generative adversarial networks (GAN) [70]. Fully convolutional network (FCN) [60] based approaches proved successful, mostly when stacking multiple simpler FCN models like the methods proposed in [73] and [66]. [52] added deep supervision [41] 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 [58] as an extension of FCN was explicitly designed for automatic semantic segmentation of biomedical data and is generally very successful. Several extensions to the U-Net architecture have been proposed. [64] combined U-Net with the Ki-Net encoder-decoder architecture to improve the segmentation of objects with blurry and noisy edges. [34] fused U-Net(+) and SegNet to address the placenta segmentation in ultrasound, where they use data similar to our work. [39] used a Multi U-Net model to perform real-time segmentation of breast ultrasound structures. [68] 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 [32] has proven successful in the broader field of biomedical data. Although it is 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.

Despite extensive research and progress in the field, the problem of automatic segmentation of the uterus in vol-

umetric ultrasound data has yet to be solved. In a recently presented approach to 3D segmentation, [6] used a 2D approach in which the model was run on each slice in all three dimensions, and results were merged to produce a 3D segmentation. This analysis loses some spatial information that can be valuable in 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. [38] 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. However, the ultrasound images and the appearance of the uterus differ significantly from our dataset because they were acquired at the later stages of pregnancy, when the uterus has a different shape and appearance, and the ultrasound examination was performed with an abdominal rather than a transvaginal ultrasound.

Recently, Zhou et al. [75] presented a study on 3D reconstruction and automatic segmentation of the uterine cavity. They used the recurrent saliency transformation network (RSTN) [67] and achieved Dice coefficient (DSC) [21, 62] of around 0.8. They were, however, using magnetic resonance imaging (MRI) data that typically provides higher resolution and less noisy images than ultrasound and can therefore be easier to segment.

1.2.2. 3D object alignment

3D data alignment is a well-researched topic with applications in numerous fields, including robotics, 3D scanning, and medical image registration. Often, classical point-based approaches are used for alignment, such as Iterative Closest Point [49, 69], Coherent Point Drift [53], and Principal Component Analysis [14]. Geometric properties such as convex hull [20] can also be used. Another option is the use of feature-based methods, which first identify salient points on 3D models, such as corners, edges, and other key points. Then, these points are linked and aligned. Examples of such methods include

Scale-Invariant Feature Transform [44], Speeded-Up Robust Features [5], and Random Sample Consensus [26]).

Many of the above methods assume that the objects to be aligned are identical or very similar. However, in some cases, alignment is required for dissimilar objects, such as aligning virtual objects with real objects in virtual and augmented reality applications or registering medical images before and after surgery. Therefore, deep learning methods [33] have gained importance in this area, as they can align different objects with a sufficient amount of data. The recently introduced ART [74] is a neural network module that can be integrated into various architectures for 3D tasks (such as point cloud auto-encoding, 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, and these are not always available.

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:

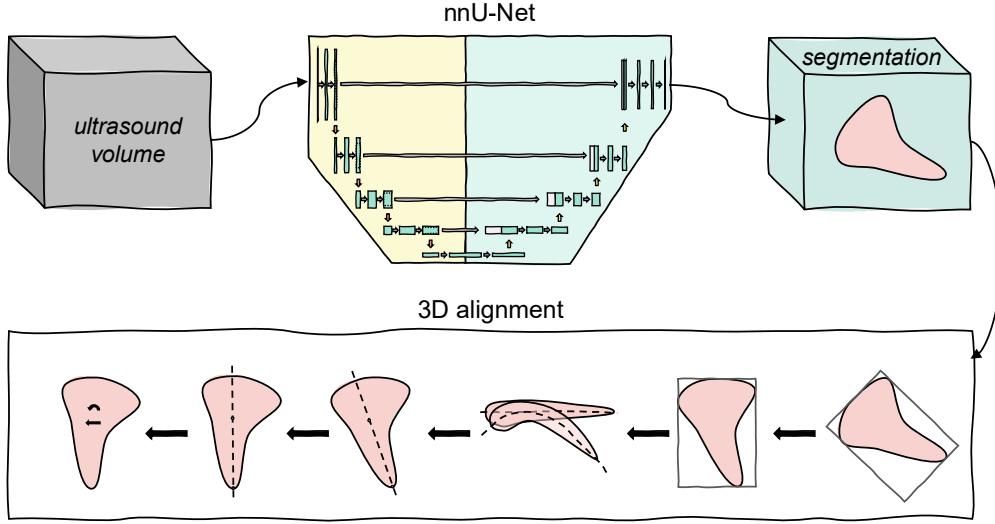


Figure 3: Diagram of the procedure from ultrasound volume to the final result - aligned uteri.

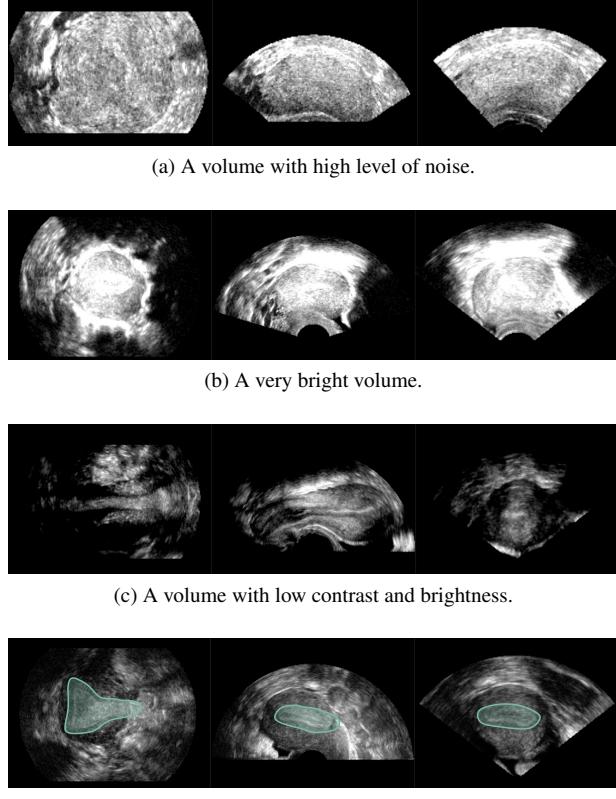
1. women who have never been pregnant aged between 18 and 28 in order to avoid acquired anomalies of the uterus that have the tendency to increase with age – *general population*,
2. women with at least 6 months of *unexplained* causes of infertility (normal semen analysis, tubal patency, no ovulation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

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 [59].

Most of our collected volumes were acquired with General Electric (GE) ultrasound equipment. The acquired ultrasound data were transformed from the original toroidal coordinate system used by the GE ultrasound machines [43] to the Cartesian coordinate system using the Slicer3D [36] extension SlicerHeart [40]. Due to the inherent physical properties of ultrasound, the imaging data does not naturally take the form of a cuboid. Therefore, during the transformation process, the resulting spherical shape is enclosed within a cuboid, with the values of the unoccupied voxels set to 0. We decided to keep the original spherical shape instead of transforming it into a cuboid (which would minimize voxel waste) in order to preserve the anatomical shapes of organs within the body. This type also proved to be more understandable when visualized by both the annotators 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 transformed 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. 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 [18] with a tool developed



(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.

Figure 4: Slices of volumes on different planes with different levels of quality.

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 Slicer3D, an open-source software for visualization, analysis, and annotation of medical, biomedical, and other 3D data. Volumes imported into Slicer3D 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 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 [48]. 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 [45]. 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 Figure 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.

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.

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 minutes 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 Figure 5.

Volumes with insufficiently visible uterine regions were

not annotated because poorly annotated data would adversely affect further processing steps. In addition, data obtained later in the research, after the segmentation model had already been trained, were not manually annotated but are included in the publicly available dataset. The dataset, named **UterUS** (Uterine UltraSound), thus contains 141 3D ultrasound volumes with manual annotations of the uterine region and 113 unannotated volumes. The volumes and their associated segmentations are stored in the NiFTi format. The dataset is available on: <https://github.com/UL-FRI-LGM/UterUS> under the CC-BY-NC-SA 4.0 licence¹.

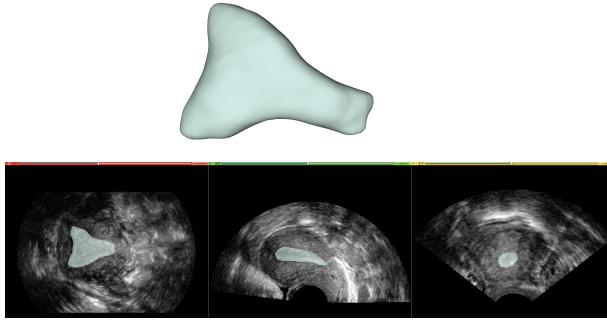


Figure 5: An annotated volume visualized with the Slicer3D tool. Above is the 3D representation of the annotated region, and below are 2D projections that allow movement along the coronal (red), axial (green), and sagittal (yellow) planes.

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 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 [72], an extension of the U-Net architecture; Med3D [16], which is pre-trained on various medical data; contrastive learning [12], which incorporates unsupervised learning and

allows for the use of unannotated data; and nnU-net [32], an extension of the U-Net [58].

The best results were obtained with the nnU-net, which we chose for all future experiments. 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 [3], the KiTS19 Challenge [30], and the BraTS Challenge [46].

We train the nnU-net model to predict the location and segmentation of the uterine cavity in 3D ultrasound volumes. The model is 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 of the model, 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 procedure in Section 3.1.

¹<https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

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 [4], normal distributions transform [7], and an algorithm based on phase correlation [22]. 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 in [74], the dataset is not big enough, while for supervised deep learning approaches, such as the method presented in [61], we lack the manual alignments.

We therefore propose a new alignment method based on geometric principles. The proposed method consists of several steps: (1) selection of the target uterus, (2) initial alignment of the uteri, (3) curvature reduction, (4) alignment of the uteri according to their principal axes, and (5) fine-tuning of the alignment.

2.3.1. Choosing the target uterus

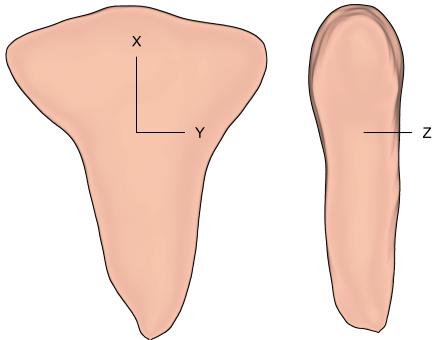


Figure 6: The uterus used for alignment from the front (left) and side (right) views with depicted axes.

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. The selected uterus shape is shown in Figure 6.

2.3.2. Initial alignment

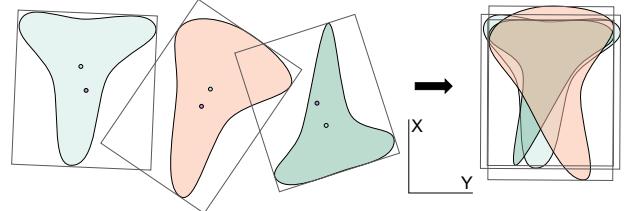


Figure 7: A scheme of the initial step of the alignment process with the OBB shown in 2D for a clearer depiction.

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) [27]. 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 (see Figure 8). 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

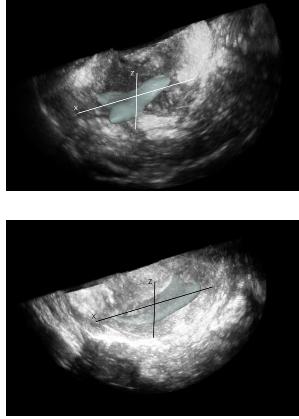


Figure 8: Two volumes rendered from the same point of view with segmented uteri rotated in different directions.

the uterus by 180 degrees around the z axis; otherwise we leave it as it is.

2.3.3. Reducing the curvature

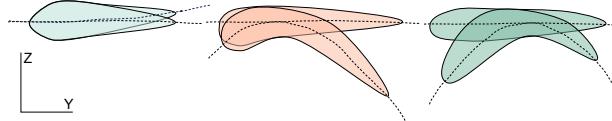


Figure 9: Three examples of the curvature reduction step in 2D. The initial surface curvature and the flattened surface are shown with dotted lines.

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 Figure 9.

To find the curvature surface, we model the uterus shape as a point cloud and look for a surface 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 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 surface:

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

where m is the number of points in the point cloud and dist is the shortest distance between a point and the surface. We use the Levenberg–Marquardt algorithm [42], 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 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

When the uteri are flattened and thus better aligned along the z axis, they must also be aligned along the other

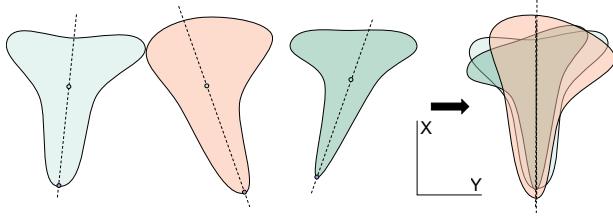


Figure 10: A schematic of the alignment step using the principal axis in 2D.

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 Figure 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.

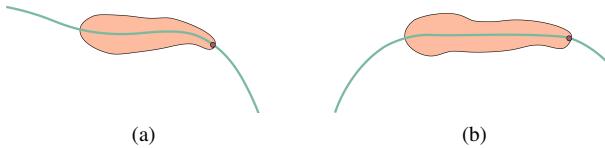


Figure 11: Uterine shapes projected onto the XZ plane with a fitted 4th order polynomial. The polynomial is shown in green and the midpoint of the cervix is red.

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. An example is shown in Figure 11. 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 centre 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

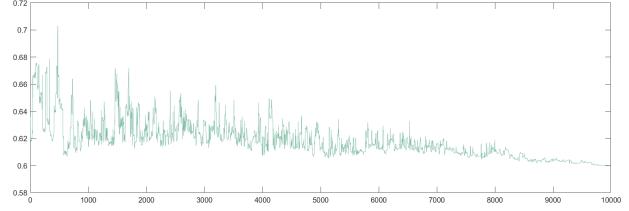


Figure 12: The process of simulated annealing, the y-axis represents energy E and the x-axis iterations.

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.

2.3.5. Fine-tuning of rotation and translation

To improve the accuracy of the alignment, we use simulated annealing [37] 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, $|\text{combined}|$ represents the number of overlapping points and $|\text{sample}|$ and $|\text{target}|$ 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

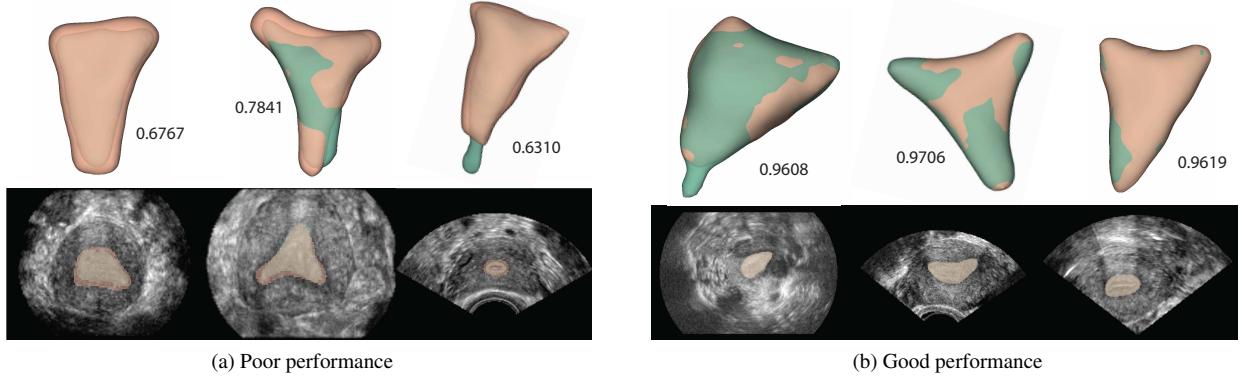


Figure 13: Volumes with the worst and best evaluation scores with DSC values. The green color represents the prediction of our model, and the orange is the ground truth from our dataset.

$[-0.35, 0.35]$ for rotation. t_N represents the *neighbour 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.

An example of how the energy decreases during the entire process is shown in Figure 12. We can see that the system changes its state more abruptly at the beginning and gradually cools down to an energy minimum.

3. Results and Discussion

3.1. Segmentation

The nnU-Net network was 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). The

Table 1: Evaluation of the nnU-Net segmentation model using the DSC metric. Means and standard deviations of five-fold models are shown for different test sets.

	Dice coefficient
Validation Set 2D	0.803 ± 0.133
Validation Set 3D	0.909 ± 0.046
Test Set (3D)	0.899 ± 0.076
Unseen Hospital Set (3D)	0.863 ± 0.043

training 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 [54] 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 [15], with the power factor set to 0.9. For the loss function, we used a compound loss function – a combination of cross-entropy loss and Dice loss [24] which has been shown to be more robust and work well on imbalanced data [47]. 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 [31].

We compared two nnU-Net models: one operating only on 2D images and a full 3D model. For out-of-domain testing, we additionally annotated 10 volumes from a hospital that was not included in the training and test-

ing datasets (the *Unseen Hospital Set*) and evaluated the model on this dataset.

The segmentation results in Table 1 show that the full-resolution 3D model performs better than the 2D model, highlighting the advantage of volumetric representation over 2D image processing. The 3D model performed well on average, with no large differences between the validation and test datasets, indicating that the model did not overfit the training data. Furthermore, the performance does not drop significantly for the out-of-domain dataset.

Looking at the three volumes with the lowest evaluation scores in Figure 13a, 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. Figure 13b 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.

3.2. Alignment

Table 2: Differences in translations and rotations of automatically and manually aligned uteri. Rotations are given in degrees, and translations are normalized to the shape length along each axis.

Translation X	Translation Y	Translation Z
3.79 ± 3.31	0.85 ± 4.13	-0.75 ± 3.49
Rotation X	Rotation Y	Rotation Z
$-3.69^\circ \pm 3.90^\circ$	$2.76^\circ \pm 5.04^\circ$	$21.80^\circ \pm 45.55^\circ$

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. Figure 14 presents a visual comparison between the automatic and manual alignments from different perspectives. Table 2 shows average differences in translations and rotations calculated between each pair of the automatically and manually aligned uteri.

We can see that our approach produced good results both visually and numerically. The visualization in Figure 14 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, numerical results in Table 2 show that translation differences are small along all axes, so automatic alignment positioned all uteri similarly. The differences in rotations along the X and Y axes are small as the initial bounding box alignment and subsequent refinements align those axes very well. A larger difference can be seen on the Z 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 de-

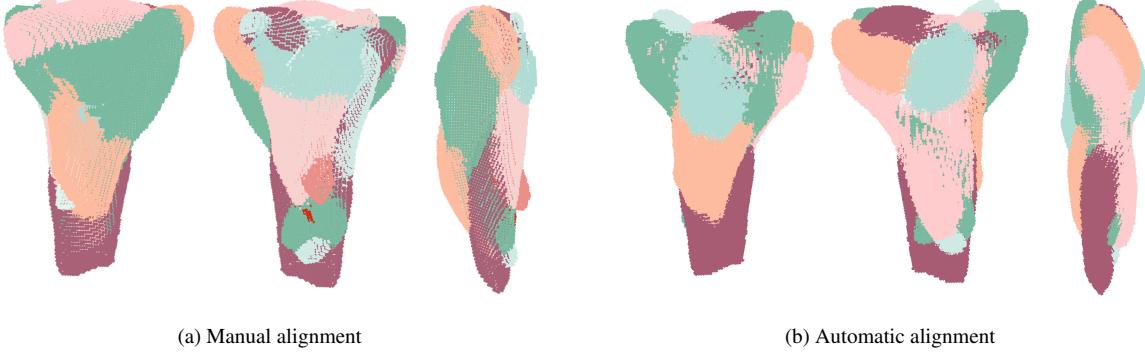


Figure 14: Visual results of the alignment process.

formations, 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. Figure 15 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 optimal. We could consider a measure that incorporates more shape information, such as graph matching of the medial scaffold as presented in [13].

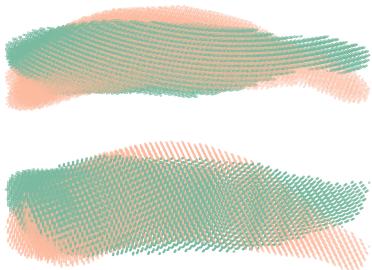


Figure 15: Uteri point clouds and their flattened shapes after curvature reduction. Original point clouds are presented with orange and flattened shapes with green color.

Alignment enables us to visualize the average 3D uterine shape, as shown in Figure 16. The visualization

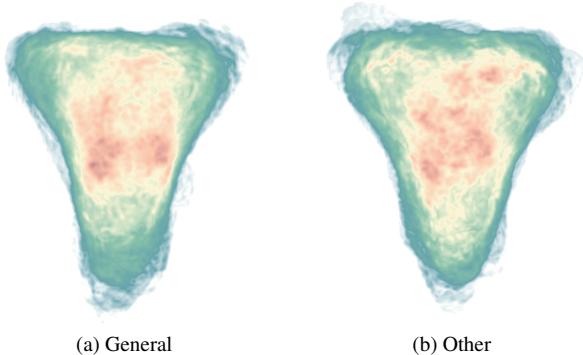


Figure 16: Visualization of the average shapes of the uterus for the general population (left) and the combined group of women with recurrent spontaneous miscarriages and unexplained infertility (right). The darker green color shows a lower density of boundary points, whereas the brighter yellow and orange present higher density.

was done with Avizo3D² and shows the average shape of the uterus for the general population (normal uteri) on the left and the average shape of the remaining two groups (women with recurrent spontaneous miscarriages and women with unexplained infertility) on the right side.

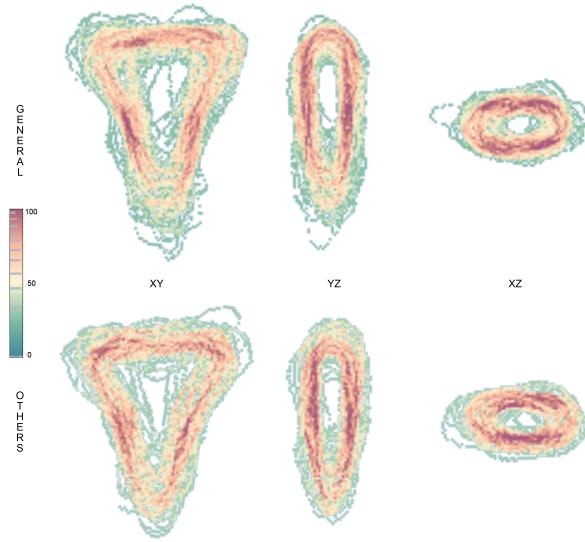


Figure 17: Standard deviations of the shape for the *general* (on the top) and *others* (on the bottom) groups shown as an octo slice from XY, YZ, and XZ orientations (left, middle and right, respectively).

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 average 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 average 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 the average shape that is very similar to

that of the general population. Figure 17 visualizes the average shape and standard deviations of automatically aligned uteri for both groups, highlighting areas of boundary clustering and deviation. 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.

Future efforts in the study will 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, enabling the true potential of our alignment method in understanding and diagnosing uterine abnormalities.

4. Conclusion

In this study, we present an approach for 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 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 out-of-domain dataset. A qualitative investigation also showed that the segmentations are accurate. 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 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 improve both the segmentation model and the final uterus reference model with a growing amount of data collected as part of the NURSE research initiative. We also plan to improve the manual annotations as currently, despite the expert review, we still find occasional errors in the manual annotations that can negatively impact the training of the segmentation model. Involving experts and multiple annotators in the annotation

²<https://www.thermofisher.com/sa/en/home/electron-microscopy/products/software-em-3d-vis/avizo-software.html>

process could improve the accuracy of the trained model, but this work is time-consuming and consequently expensive. In addition to the future improvements of the alignment method already mentioned in section 3.2, the larger dataset will allow us to use deep learning alignment methods that may prove to be more robust than the proposed approach.

We hope that this work will lead to new insights in the field of gynecology, and the initial feedback from participants in the NURSE multi-centre study has been very positive. The preliminary results presented at the 32nd Annual Meeting of the ESGE (European Society for Gynaecological 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, which could shorten the time to conception and reduce the rate of spontaneous miscarriages.

References

- [1] Abu-Zidan, F.M., Hefny, A.F., Corr, P., 2011. Clinical ultrasound physics. *Journal of Emergencies, Trauma and Shock* 4. doi:10.4103/0974-2700.86646.
- [2] Ahn, H., Hernández-Andrade, E., Romero, R., Ptwardhan, M., Goncalves, L.F., Aurioles-Garibay, A., Garcia, M., Hassan, S.S., Yeo, L., 2013. Mirror artifacts in obstetric ultrasound: Case presentation of a ghost twin during the second-trimester ultrasound scan. *Fetal Diagnosis and Therapy* 34. doi:10.1159/000353702.
- [3] Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., van Ginneken, B., Bilello, M., Bilic, P., Christ, P.F., Do, R.K., Gollub, M.J., Heckers, S.H., Huisman, H., Jaragnin, W.R., McHugo, M.K., Napel, S., Pernicka, J.S., Rhode, K., Tobon-Gomez, C., Vorontsov, E., Meakin, J.A., Ourselin, S., Wiesenfarth, M., Arbeláez, P., Bae, B., Chen, S., Daza, L., Feng, J., He, B., Isensee, F., Ji, Y., Jia, F., Kim, I., Maier-Hein, K., Merhof, D., Pai, A., Park, B., Perslev, M., Rezaifar, R., Rippel, O., Sarasua, I., Shen, W., Son, J., Wachinger, C., Wang, L., Wang, Y., Xia, Y., Xu, D., Xu, Z., Zheng, Y., Simpson, A.L., Maier-Hein, L., Cardoso, M.J., 2022. The Medical Segmentation Decathlon. *Nature Communications* 13. doi:10.1038/s41467-022-30695-9.
- [4] Arun, K.S., Huang, T.S., Blostein, S.D., 1987. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9*. doi:10.1109/TPAMI.1987.4767965.
- [5] Bay, H., Tuytelaars, T., Van Gool, L., 2006. SURF: Speeded up robust features, in: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/11744023__32.
- [6] Behboodi, B., Rivaz, H., Lalondrelle, S., Harris, E., 2021. Automatic 3D Ultrasound Segmentation of Uterus Using Deep Learning, in: *IEEE International Ultrasonics Symposium, IUS*. doi:10.1109/IUS52206.2021.9593671.
- [7] Biber, P., 2003. The Normal Distributions Transform: A New Approach to Laser Scan Matching, in: *IEEE International Conference on Intelligent Robots and Systems*. doi:10.1109/iros.2003.1249285.
- [8] Biswas, M., Kuppili, V., Edla, D.R., Suri, H.S., Saba, L., Marinhoe, R.T., Sanches, J.M., Suri, J.S., 2018. Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm. *Computer Methods and Programs in Biomedicine* 155. doi:10.1016/j.cmpb.2017.12.016.
- [9] Bolliger, C., Herth, F., Mayo, P., Miyazawa, T., Beamis, J., 2009. Clinical chest ultrasound: from the ICU to the bronchoscopy suite.
- [10] Bridal, S.L., Correas, J.M., Saïed, A., Laugier, P., 2003. Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging, in: *Proceedings of the IEEE*. doi:10.1109/JPROC.2003.817879.

- [11] Cao, Z., Duan, L., Yang, G., Yue, T., Chen, Q., 2019. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Medical Imaging* 19. doi:10.1186/s12880-019-0349-x.
- [12] Chaitanya, K., Erdil, E., Karani, N., Konukoglu, E., 2020. Contrastive learning of global and local features for medical image segmentation with limited annotations, in: *Advances in Neural Information Processing Systems*.
- [13] Chang, M.C., Kimia, B.B., 2011. Measuring 3D shape similarity by graph-based matching of the medial scaffolds. *Computer Vision and Image Understanding* 115. doi:10.1016/j.cviu.2010.10.013.
- [14] Chaouch, M., Verroust-Blondet, A., 2009. Alignment of 3D models. *Graphical Models* 71. doi:10.1016/j.gmod.2008.12.006.
- [15] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40. doi:10.1109/TPAMI.2017.2699184.
- [16] Chen, S., Ma, K., Zheng, Y., 2019. Med3d: Transfer learning for 3d medical image analysis. *CoRR* abs/1904.00625.
- [17] Chi, J., Walia, E., Babyn, P., Wang, J., Groot, G., Eramian, M., 2017. Thyroid Nodule Classification in Ultrasound Images by Fine-Tuning Deep Convolutional Neural Network. *Journal of Digital Imaging* 30. doi:10.1007/s10278-017-9997-y.
- [18] Cox, R.W., Ashburner, J., Breman, H., Fissell, K., Haselgrave, C., Holmes, C.J., Lancaster, J.L., Rex, D.E., Smith, S.M., Woodward, J.B., Strother, S.C., 2004. A (Sort of) New Image Data Format Standard: NIfTI-1. *NeuroImage* 22.
- [19] Cunningham, R., Sánchez, M., Loram, I., 2019. Ultrasound segmentation of cervical muscle during head motion: A dataset and a benchmark using deconvolutional neural networks. *Engineering Engrxiv Archive* .
- [20] Cupec, R., Vidović, I., Filko, D., Djurović, P., 2020. Object recognition based on convex hull alignment. *Pattern Recognition* 102. doi:10.1016/j.patcog.2020.107199.
- [21] Dice, L.R., 1945. Measures of the Amount of Ecologic Association Between Species. *Ecology* 26. doi:10.2307/1932409.
- [22] Dimitrievski, M., Van Hamme, D., Veelaert, P., Philips, W., 2016. Robust Matching of Occupancy Maps for Odometry in Autonomous Vehicles. doi:10.5220/0005719006260633.
- [23] Driscoll, P., 2006. Gray's Anatomy, 39th Edition. *Emergency Medicine Journal* 23. doi:10.1136/emj.2005.027847.
- [24] Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of skip connections in biomedical image segmentation, in: *Lecture Notes in Computer Science (including sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-319-46976-8__19.
- [25] Ellis, H., 2005. Anatomy of the uterus. *Anaesthesia and Intensive Care Medicine* 6. doi:10.1383/anes.6.3.74.62222.
- [26] Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24. doi:10.1145/358669.358692.
- [27] Gottschalk, S., Manocha, D., Lin, M.C., 2000. Collision Queries using Oriented Bounding Boxes. Phd thesis. The University of North Carolina at Chapel Hill.
- [28] Grimbizis, G.F., Gordts, S., Di Spiezio Sardo, A., Brucker, S., De Angelis, C., Gergolet, M., Li, T.C., Tanos, V., Brölmann, H., Gianaroli, L., Campo, R.,

2013. The ESHRE/ESGE consensus on the classification of female genital tract congenital anomalies. *Human Reproduction* 28. doi:10.1093/humrep/det098.
- [29] Harrison, P.J., Freemantle, N., Geddes, J.R., 2003. Meta-analysis of brain weight in schizophrenia. *Schizophrenia Research* 64. doi:10.1016/S0920-9964(02)00502-9.
- [30] Heller, N., Isensee, F., Maier-Hein, K.H., Hou, X., Xie, C., Li, F., Nan, Y., Mu, G., Lin, Z., Han, M., Yao, G., Gao, Y., Zhang, Y., Wang, Y., Hou, F., Yang, J., Xiong, G., Tian, J., Zhong, C., Ma, J., Rickman, J., Dean, J., Stai, B., Tejpaul, R., Oestreich, M., Blake, P., Kaluzniak, H., Raza, S., Rosenberg, J., Moore, K., Walczak, E., Rengel, Z., Edgerton, Z., Vasdev, R., Peterson, M., McSweeney, S., Peterson, S., Kalapara, A., Sathianathan, N., Pananikopoulos, N., Weight, C., 2021. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge. *Medical Image Analysis* 67. doi:10.1016/j.media.2020.101821.
- [31] HPC RIVR consortium, 2022. HPC System Vega. Available at www.hpc-rivr.si.
- [32] Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* 18. doi:10.1038/s41592-020-01008-z.
- [33] Jaderberg, M., Simonyan, K., Zisserman, A., Kavukcuoglu, K., 2015. Spatial transformer networks, in: *Advances in Neural Information Processing Systems*.
- [34] Jain, P.K., Sharma, N., Giannopoulos, A.A., Saba, L., Nicolaides, A., Suri, J.S., 2021. Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound. *Computers in Biology and Medicine* 136. doi:10.1016/j.combiomed.2021.104721.
- [35] Joel, T., Sivakumar, R., 2018. An extensive review on Despeckling of medical ultrasound images using various transformation techniques. *Applied Acoustics* 138. doi:10.1016/j.apacoust.2018.03.023.
- [36] Kikinis, R., Pieper, S.D., Vosburgh, K.G., 2014. 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support, in: *Intraoperative Imaging and Image-Guided Therapy*. doi:10.1007/978-1-4614-7657-3\}19.
- [37] Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* 220. doi:10.1126/science.220.4598.671.
- [38] Kong, B., Wang, X., Lu, Y., Yang, H.Y., Cao, K., Song, Q., Yin, Y., 2022. Ultrasound Uterus and Fetus Segmentation with Constrained Spatial-Temporal Memory FCN, in: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi:10.1007/978-3-031-12053-4\}19.
- [39] Kumar, V., Webb, J.M., Gregory, A., Denis, M., Meixner, D.D., Bayat, M., Whaley, D.H., Fatemi, M., Alizad, A., 2018. Automated and real-time segmentation of suspicious breast masses using convolutional neural network. *PLoS ONE* 13. doi:10.1371/journal.pone.0195816.
- [40] Lasso, A., Herz, C., Nam, H., Cianciulli, A., Pieper, S., Drouin, S., Pinter, C., St-Onge, S., Vigil, C., Ching, S., Sunderland, K., Fichtinger, G., Kikinis, R., Jolley, M.A., 2022. SlicerHeart: An open-source computing platform for cardiac image analysis and modeling. doi:10.3389/fcvm.2022.886549.
- [41] Lee, C.Y., Xie, S., Gallagher, P.W., Zhang, Z., Tu, Z., 2015. Deeply-supervised nets, in: *Journal of Machine Learning Research*.
- [42] Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* 2. doi:10.1090/qam/10666.
- [43] Looney, P., Stevenson, G., Collins, S., 2019. 3D ultrasound file reading and coordinate transformations. *Journal of Open Source Software* 4. doi:10.21105/joss.01063.

- [44] Lowe, D.G., 1999. Object recognition from local scale-invariant features, in: Proceedings of the IEEE International Conference on Computer Vision. doi:10.1109/iccv.1999.790410.
- [45] Lucidi, T., 2021. Noise artifacts in diagnostic ultrasound - innovatus imaging. <https://www.innovatusimaging.com/whats-that-noise-part-1/>.
- [46] Luu, H.M., Park, S.H., 2022. Extending nn-UNet for Brain Tumor Segmentation, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi:10.1007/978-3-031-09002-8{_}16.
- [47] Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., Martel, A.L., 2021. Loss odyssey in medical image segmentation. Medical Image Analysis 71. doi:10.1016/j.media.2021.102035.
- [48] Maar, M., Lee, J., Tardi, A., Zheng, Y.Y., Wong, C., Gao, J., 2022. Inter-transducer variability of ultrasound image quality in obese adults: Qualitative and quantitative comparisons. Clinical Imaging 92. doi:10.1016/j.clinimag.2022.09.010.
- [49] Mäkelä, T., Clarysse, P., Sipilä, O., Pauna, N., Pham, Q.C., Katila, T., Magnin, I.E., 2002. A review of cardiac image registration methods. doi:10.1109/TMI.2002.804441.
- [50] Menchón-Lara, R.M., Sancho-Gómez, J.L., Bueno-Crespo, A., 2016. Early-stage atherosclerosis detection using deep learning over carotid ultrasound images. Applied Soft Computing Journal 49. doi:10.1016/j.asoc.2016.08.055.
- [51] Middleton, W.D., 2012. Ultrasound the requisites.
- [52] Mishra, D., Chaudhury, S., Sarkar, M., Soin, A.S., 2019. Ultrasound Image Segmentation: A Deeply Supervised Network With Attention to Boundaries. IEEE Transactions on Biomedical Engineering 66. doi:10.1109/TBME.2018.2877577.
- [53] Myronenko, A., Song, X., Carreira-Perpiñán, M., 2006. Non-rigid point set registration: Coherent Point Drift, in: NIPS 2006: Proceedings of the 19th International Conference on Neural Information Processing Systems. doi:10.7551/mitpress/7503.003.0131.
- [54] Nesterov, Y., 1983. A method of solving a convex programming problem with convergence rate $\$O(1/k^2)\$$. Soviet Mathematics Doklady 27.
- [55] Nott, J., 2020. Anatomy of the uterus — female reproductive anatomy — geeky medics. <https://geekymedics.com/anatomy-of-the-uterus/>.
- [56] Oppelt, P., Renner, S.P., Brucker, S., Strissel, P.L., Strick, R., Oppelt, P.G., Doerr, H.G., Schott, G.E., Hucke, J., Wallwiener, D., Beckmann, M.W., 2005. The VCUAM (Vagina Cervix Uterus Adnex-associated Malformation) classification: A new classification for genital malformations. Fertility and Sterility 84. doi:10.1016/j.fertnstert.2005.05.036.
- [57] Pu, B., Li, K., Li, S., Zhu, N., 2021. Automatic Fetal Ultrasound Standard Plane Recognition Based on Deep Learning and IIoT. IEEE Transactions on Industrial Informatics 17. doi:10.1109/TII.2021.3069470.
- [58] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi:10.1007/978-3-319-24574-4{_}28.
- [59] Saravelos, S.H., Li, T.C., 2016. Intra-cycle variation of the uterine cavity indentation assessed with three-dimensional ultrasound in natural and stimulated cycles. Reproductive BioMedicine Online 32. doi:10.1016/j.rbmo.2016.02.005.
- [60] Shelhamer, E., Long, J., Darrell, T., 2017. Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39. doi:10.1109/TPAMI.2016.2572683.
- [61] Shimada, S., Golyanik, V., Treitschk, E., Stricker, D., Theobalt, C., 2019. DispVoxNets: Non-Rigid Point

- Set Alignment with Supervised Learning Proxies, in: Proceedings - 2019 International Conference on 3D Vision, 3DV 2019. doi:10.1109/3DV.2019.00013.
- [62] Sørensen, T., 1948. 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. Biological Skr 5.
- [63] Tutschek, B., 2017. Three-dimensional ultrasound: Techniques and clinical applications, in: Obstetric Imaging: Fetal Diagnosis and Care, 2nd Edition. doi:10.1016/B978-0-323-44548-1.00173-X.
- [64] Valanarasu, J.M.J., Sindagi, V.A., Hacihaliloglu, I., Patel, V.M., 2020. KiU-Net: Towards Accurate Segmentation of Biomedical Images Using Over-Complete Representations, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi:10.1007/978-3-030-59719-1__36.
- [65] Wolf, D.C., 1990. Evaluation of the Size, Shape, and Consistency of the Liver.
- [66] Wu, L., Xin, Y., Li, S., Wang, T., Heng, P.A., Ni, D., 2017. Cascaded Fully Convolutional Networks for automatic prenatal ultrasound image segmentation, in: Proceedings - International Symposium on Biomedical Imaging. doi:10.1109/ISBI.2017.7950607.
- [67] Xie, L., Yu, Q., Zhou, Y., Wang, Y., Fishman, E.K., Yuille, A.L., 2020. Recurrent Saliency Transformation Network for Tiny Target Segmentation in Abdominal CT Scans. IEEE Transactions on Medical Imaging 39, 514–525. doi:10.1109/TMI.2019.2930679.
- [68] Xu, L., Gao, S., Shi, L., Wei, B., Liu, X., Zhang, J., He, Y., 2021. Exploiting vector attention and context prior for ultrasound image segmentation. Neurocomputing 454. doi:10.1016/j.neucom.2021.05.033.
- [69] Yang, J., Li, H., Campbell, D., Jia, Y., 2016. Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration. IEEE Transactions on Pattern Analysis and Machine Intelligence 38. doi:10.1109/TPAMI.2015.2513405.
- [70] Yang, X., Dou, H., Li, R., Wang, X., Bian, C., Li, S., Ni, D., Heng, P.A., 2018. Generalizing Deep Models for Ultrasound Image Segmentation, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi:10.1007/978-3-030-00937-3__57.
- [71] Yang, X., Yu, L., Li, S., Wen, H., Luo, D., Bian, C., Qin, J., Ni, D., Heng, P.A., 2019. Towards automated semantic segmentation in prenatal volumetric ultrasound. IEEE Transactions on Medical Imaging 38. doi:10.1109/TMI.2018.2858779.
- [72] Žerovnik Mekuč, M., Bohak, C., Hudoklin, S., Kim, B.H., Romih, R., Kim, M.Y., Marolt, M., 2020. Automatic segmentation of mitochondria and endolysosomes in volumetric electron microscopy data. Computers in Biology and Medicine 119. doi:10.1016/j.combiomed.2020.103693.
- [73] Zhang, Y., Ying, M.T., Yang, L., Ahuja, A.T., Chen, D.Z., 2017. Coarse-to-Fine Stacked Fully Convolutional Nets for lymph node segmentation in ultrasound images, in: Proceedings - 2016 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2016. doi:10.1109/BIBM.2016.7822557.
- [74] Zhou, K., Bhatnagar, B.L., Schiele, B., Pons-Moll, G., 2022. Adjoint Rigid Transform Network: Task-conditioned Alignment of 3D Shapes, in: Proceedings - 2022 International Conference on 3D Vision, 3DV 2022. doi:10.1109/3DV57658.2022.00019.
- [75] Zhou, Q., Bai, Y., Chen, F., Zhang, H., Chen, L., Zhang, G., Wang, Y., Sui, L., 2024. 3D reconstruction of the uterus and automatic segmentation of the uterine cavity on 3D magnetic resonance imaging: A preliminary study. Heliyon 10, e23558. URL: <https://www.sciencedirect.com/science/>

article/pii/S2405844023107663, doi:<https://doi.org/10.1016/j.heliyon.2023.e23558>.

- [76] Zhou, R., Fenster, A., Xia, Y., Spence, J.D., Ding, M., 2019. Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images. *Medical Physics* 46. doi:10.1002/mp.13581.