Low-Limb Muscles segmentation in 3D Freehand Ultrasound using Non-Learning Methods and Label Transfer.

Vanessa Gonzalez Duque[a], Dawood Al Chanti[a], Marion Crouzier[b], Antoine Nordez[b], Lilian Lacourpaille[b], and Diana Mateus[a]

[a]École Centrale de Nantes, Laboratoire des Sciences du Numérique de Nantes LS2N, UMR CNRS 6004 Nantes, France.
[b]Université de Nantes, Laboratoire "Movement - Interactions - Performance", MIP, EA 4334, F-44000 Nantes, France.

ABSTRACT

In this paper, we aim to assist the measurement of the volume of lower-limb muscles from 3D freehand ultrasound images. Volume estimation typically requires a very time-consuming manual segmentation step. To facilitate and speed-up the volume measurements, in this paper, we propose a non-learning based approach that, starting from sparse annotations of muscles in 2D images, propagates the labels to the full volume. Furthermore, we rely on 3D-3D image-based registration to combine labels from different ultrasound acquisitions of the same muscle with different qualities. Our goal here is to provide a simple and low-cost solution, relying mainly on open-source software for the processing steps. The proposed approach effectively reduces the manual interaction time while providing reasonable estimations for the segmentation and volume calculation. We achieve a mean dice score of 0.89 ± 0.03 and a mean volumetric measure error of 4,18%. The resultant volumes may also be useful for building augmented annotated databases to develop automatic learning-based segmentation approaches.

Keywords: 3D freehand Ultrasound, label propagation, volume measurements

1. INTRODUCTION

Each year, one over 3500 children suffers from Duchenne Muscular Dystrophy (DMD). This degenerative muscular disorder affects men, generating progressive muscle atrophy and loss of autonomy at an early age. The amount of low-limb muscle's volume is an important discriminative metric for diagnosis and follow-up.[1] It is also considered when validating new therapies.[2] Volumes are typically measured from Computer Tomography (CT) and Magnetic Resonance (MR) images. These modalities are, however, impractical for young children. Therefore, recently Pichiecchio *et.al.*[3] has pointed to US imaging instead.

US has the advantage of being a fully noninvasive and risk-free procedure while permitting a real-time acquisition. It is also relatively inexpensive compared to MRI (the gold standard for volume calculation). However, the image quality of US is inferior, making its interpretation harder. In particular, the nature of ultrasound physics results in images with low signal-to-noise ratio, speckle, attenuation, and missing borders. Furthermore, the contrast between areas of interest is often low.

In practice, to measure the volume occupied by a muscle from an ultrasound image, an expert needs to delineate the muscle borders on 2D slices. Delineation could be done either on partial but high-resolution B-mode images, or on 2-D slices extracted from a 3D image after reconstruction (compounding of an optically tracked freehand acquisition). However, The complete delineation of all the slices in an acquisition/volume is too time-consuming and impractical when dealing with several patients.

The first objective of this work is to propose a simple yet effective semi-automatic method to alleviate the delineation efforts of the expert. We favor methods that reduce the manual segmentation time while accurately

Further author information: (Send correspondence to Vanessa Gonzalez Duque)
Vanessa Gonzalez Duque: E-mail: vanessa.gonzalez-duque@ls2n.fr

approximating the volume estimation. This step is also crucial for creating reliable "ground truth" segmentation masks,[4] which are the main requirement for later developing any machine learning algorithm.

The second objective of this study is to investigate the feasibility of label transfer between acquisitions acquired under different parameter settings. Such acquisitions are required since certain anatomical structures are deeper than others, and thus better seen under lower frequencies.[5] Experts will segment the structures only on the acquisition where the muscle is the most visible, even if present in both. Such selective segmentation generates volumes where only some of the muscles are annotated. We propose here to exploit 3D image registration, to transfer the manual annotations performed on the most visible sequence to the other acquisition where the muscle in question is less visible but still present. Thereby, we offer complete anatomical labels over each acquisition, which can improve the variability of a database for training, and thus the generalization ability of a learning-based automatic segmentation algorithm.[6]

In our previous work ,[6] we have proposed a learning-based segmentation approached trained only on incomplete masks from low-resolution 2D images obtained after 3D compounding. In this paper, we improve the quality of the training dataset at a low computational cost, by first, propagating annotations done on 2D B-mode images before the reconstruction step, which have a higher resolution; and second, by transferring the annotations between two acquisitions of the same patients.

Our contributions are:

1. We demonstrate that segmentation masks made on acquisitions with different parameters can be transferred to each other, by means of 3D spatial registration.

2. We propose a semi-automatic segmentation method, useful to accelerate the annotation of 3D US freehand acquisitions. The method starts from the partial annotations of muscles on a selection high-resolution B-mode images. Based on the image pose we are able to propagate the annotations first along each sweep before 3D reconstruction. The approach leads to a mean dice of 89% on a database of 10 subjects with volumes of size $564x632x1443 \pm (49x38x207)$.

2. RELATED WORK

US image segmentation is a challenging task due to the intrinsic characteristics of the modality. Although there is currently a plethora of learning based US segmentation methods ,[4] we are interested here in the upstream procedure to build good databases for training with minimal expert efforts. Despite being crucial, such approaches are less documented nowadays. Yoshizumi *et.al.*[7] propose a multiple-frequency ultrasonic imaging approach to enhance the image resolution by superimposing and mixing several images with different frequencies. In this way, more structures become visible facilitating their segmentation. Although effective, such a method modifies the image content itself. Here we also make acquisitions with two frequencies, but we use the most adapted one for the manual segmentation of each muscle. For transferring annotations across the two sequences we rely on mono-modal image-based rigid registration .[8]

When done manually ultrasound-segmentation, besides being difficult, is very time-consuming. The aim of this work is not to develop a novel and general segmentation algorithm, but instead to propose a simple approach to accelerate the manual segmentation process in the context of freehand ultrasound acquisitions. Several software solutions propose interactive or semi-automatic segmentation tools to propagate initial annotations or seeds. However, clinical experts are still required to initialize the seeds and refine the segmentation after automatic propagation. Software such as Slicer,[9] Imfusion[10] and Stradwin,[11] offer non-learning based built-in functions such as "Fill Between Slices", "grow from seeds" and "watershed" for label propagation.

Such non-learning based propagation methods can be divided into two categories. The first does propagation over binary label masks only, while the second takes into account the image content on top of the label-seeds. In the category of label-interpolations approaches, the "Fill Between Slices" method [12] from Slicer is an iterative morphological contour interpolator method using pixel morphological dilatation to create a gradual change in the object shape. Similarly, the "maximal disc-guided interpolation" method[13] from Stradwin interpolates a surface through sparse non-parallel labels. Both methods[12,12] perform interpolation on the 3D volume after

reconstruction. Instead, we propose here to interpolate between two partial expert annotations done on 2D B-mode images along the sweep to achieve higher resolution masks.

In the second category are methods that take into account the image content. In addition to enforcing a smooth transition between the annotated slices, they push the segmented region's boundaries to coincide with the image contours. Here we find watershed [14] as well as graph-cut [15] methods. Most of the built-in approaches assume homogeneous areas of interest and well-defined image contours. Despite the competitive performance in other modalities, the assumptions above do not hold in the case of muscles in US images, resulting in leakage. Without a specialized modification, the simple built-in implementation requires a large amount of labeled background and foreground seeds.

In this paper, we combine a binary label propagation directly on the B-mode images, with the label transfer across acquisitions of two-frequencies. This simple solution provides a non-learning-based baseline for the segmentation of volumes in 3D freehand ultrasound images while alleviating the annotation burden when building databases for supervised segmentation methods.

3. MATERIALS AND METHODS

3.1 Dataset Description

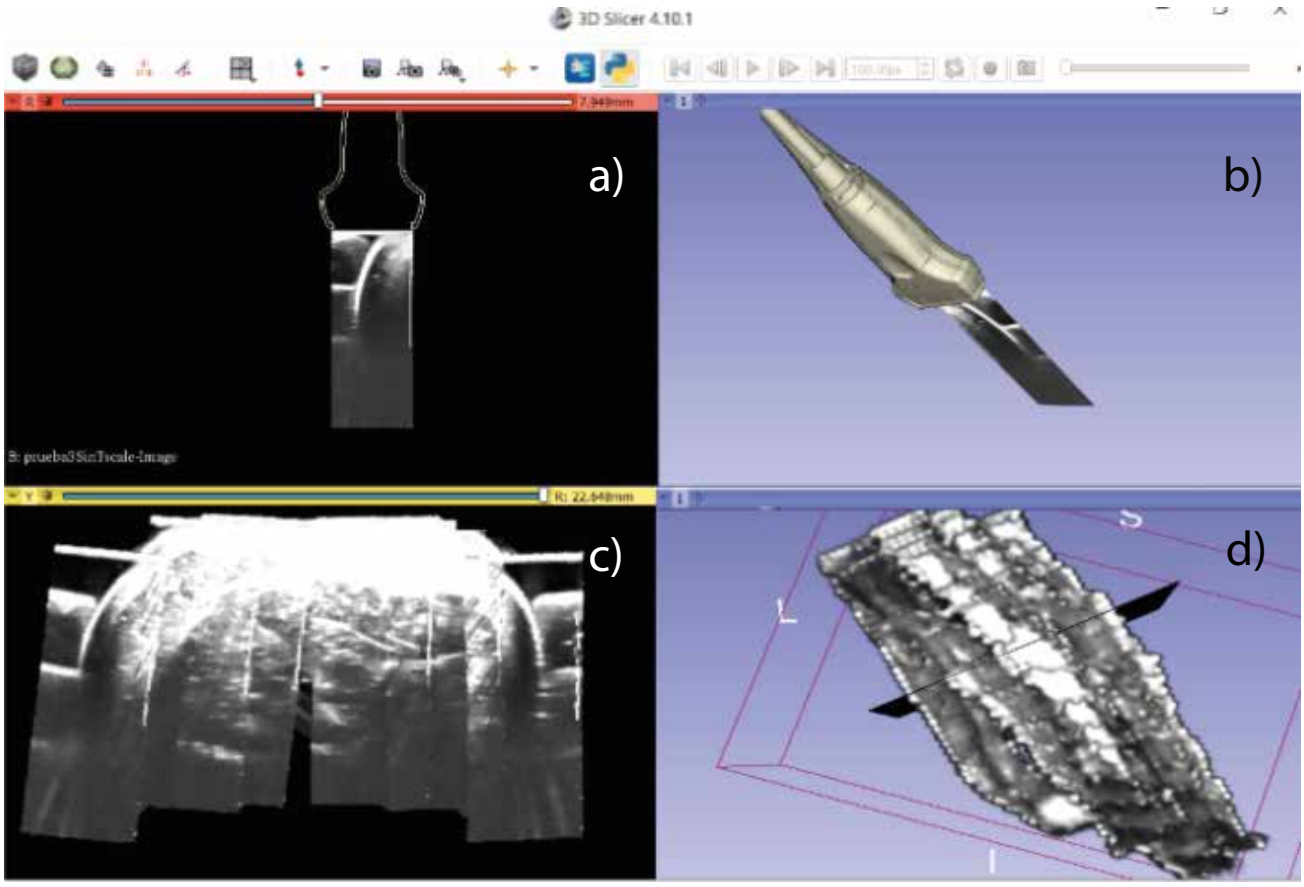

Figure 1. Volume compounding using 3D slicer. a) B-mode image. b)Spatial displacement of US images in low speed. c)Sagittal cross-section of the 3D volume. d) 3d compounded US volume of the lower leg.

Our dataset comes from the study of Crouzier et al.[16] Each participant stays prone, with their lower leg in a custom-made water bath while 2 US freehand image sequences with different parameters were recorded (See Fig 2). Four to six parallel sweeps were performed from the knee to the ankle, and the probe was tracked

with reflected points optically*. B-mode images (3.9 cm width, 9.5 cm depth±1.2) were recorded each 5 mm displacement in low-speed †. 3D ultrasound volumes were compounded using the tracking matrices of the probe, filling a voxel grid of $564 \times 632 \times 1443 \pm (49 \times 38 \times 207)$. With a pixel spacing of 0.276993mm/pixel ± 0.015. See fig 1.

Out of the original 44 participants we select 10 with 2 recordings (x_1 and x_2). For this selection, we have both the annotation of a selection of the 2D B-mode images that will be used as seeds, but also the manual annotation of every slice in the reconstructed 3D volume for the quantitative evaluation.

Image-based **Rigid Registration** is used to align the 2 reconstructed US volumes of the same participant bringing similar structures to the same spatial position in a common grid. This registration was done with imfusion software,[10] but it can be done with other similar tools. After registration, the muscle structures overlap and segmentation can be done indistinctly over one register or the other.

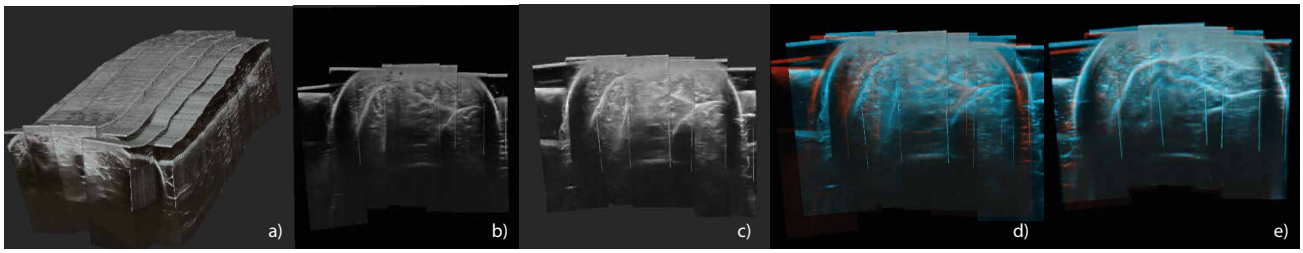

Figure 2. a)5 B-mode image sweeps. b)First acquisition c)Second acquisition d)Misaligned volumes e)Volumes after registration.

3.2 Manual Segmentation

Slice by slice segmentation over the cross-sectional area (CSA) is proved to be the most reliable method to perform skeletal volume muscle quantification[4] over MRI images. Placed on the "sagittal" plane, the annotator scrolls up and down for each structure to have an idea of how muscles evolve in order to relate them to anatomical knowledge. The following recommendations were followed to harmonize the masks from different annotators: Identify the brightest structures as bones (fibula and tibia) and ligaments. Use the knowledge that muscles separate from the bone at two-thirds of the leg. First, proceed with the annotations over slices with high certainty, with a 3D spherical brush. When possible, provide smooth annotations with slices ahead and backwards. In case of unclear boundaries, use the shape information from other anatomical structures in the nearby context. Tracing those structures until the problematic slice to decrease ambiguities. The training on the software demands around 5 hours of adaptation, and the segmentation of 3 muscles in 10 patients takes around 25 hours, equivalent to 149 minutes per patient, and 50 per muscle.

3.3 Proposed method

In this section, we describe the label propagation method to semi-automatically segment the full volume starting from sparse seeds. We then describe the method to transfer annotations from one sequence to a second one acquired with a different ultrasound frequency, in order to provide fully annotated volumes.

First, partial annotations on some of the high-resolution B-mode images are collected (Fig 4 a and b). Having two such masks over B-mode images on a sweep (Fig 4 b), the zero-order label propagation consists of copying the smallest mask to all unlabelled images in between. This is done for every pair of available partial masks. An intermediate step during this "sweep-interpolation" process is shown in Fig 4 c). We then apply a smoothing

*Optitrack system, Natural point, USA

†Recorded with a 40mm linear VERMON probe: 2-10 MHz; and Aixplorer, Supersonic Imagine Ultrasound machine, Aix-en-Provence, France

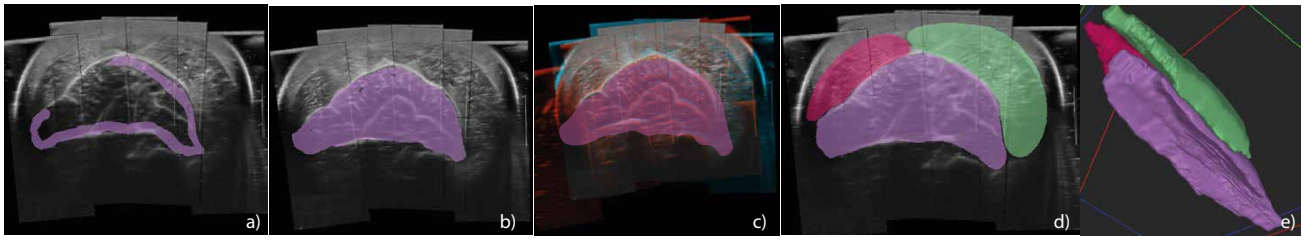

Figure 3. Manually annotated ("ground truth") dataset: a)Inferior borders delineation with spherical brush b) Muscle contour c)Annotations overlap d)Masks over hard register d)3D muscles segmentation.

Gaussian filter of kernel size 7x7 to the compounded interpolated volume (Fig 4d), which leads to a smooth muscle mask (Fig 4 e). Fig 4 f) presents an example of the obtained propagation on the sagittal plane, closely following the edges of the Soleus structure.

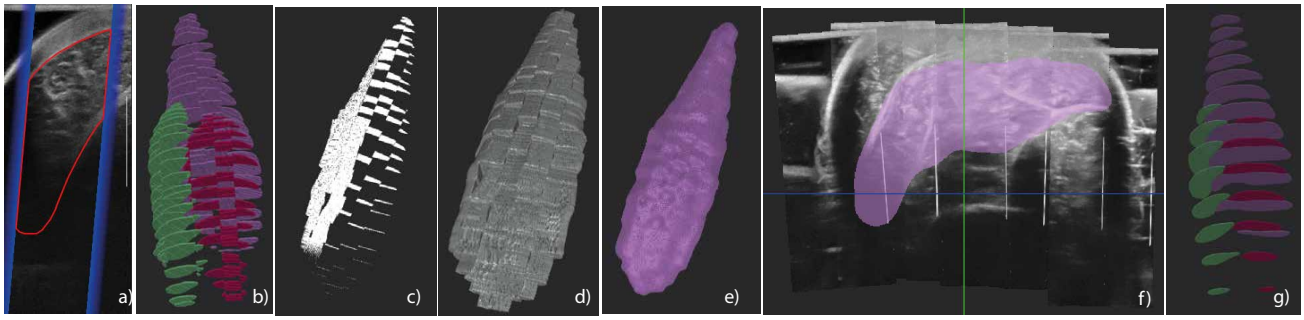

Figure 4. a) B-mode image and partial mask in red. b) Initial partial annotations used as seeds for the label propagation method. c) Zero order interpolation method in process d) Volume compounding from interpolation e) Volume compounding after smoothness. f) Example of a propagated mask after compounding and smoothness for the Soleus muscle. The purple region is a cross-sectional slice of the volume g) Equivalent amount of parallel annotations.

Gastrocnemius medial (GM) and Gastrocnemius lateral (GL) segmentations were done over a first acquisition, called x_1. Soleus (SOL) segmentation was done over the second one with more gain and less frequency, called x_2. For each participant, with the two registers x_1 and x_2, we apply monomodal image-based rigid registration algorithm.

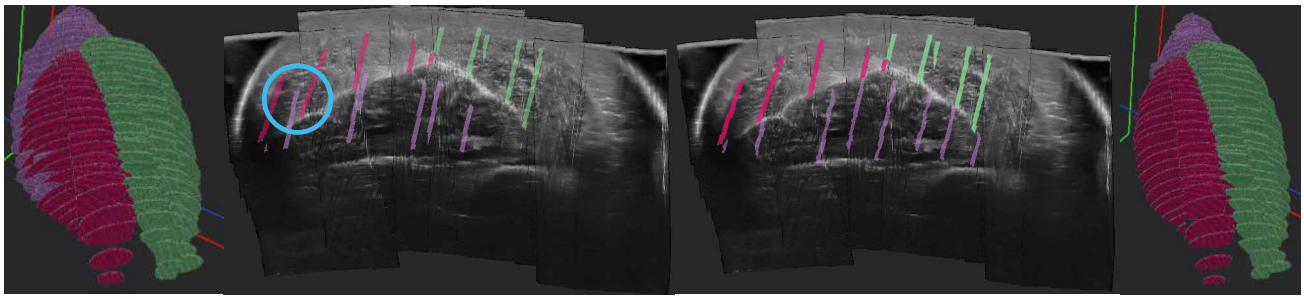

Figure 5. left) sparse annotations from the two acquisitions before alignment. The GM (green) and GL (red) masks were annotated on the first acquisition, while the Soleous (purple) masks were annotated on the second. Middle-left) Transverse view of the misaligned sparse annotations overlaid on the US volume. Middle-right) Transverse view of the sparse annotations overlaid on the US volume after registration. Right) Aligned sparse annotations.

Decoupling the annotation of organs at different depths and transferring the annotations from the best-suited acquisition to the less suited one improves the quality of the masks.

The overall proposed approach consists of several simple steps making it highly reproducible while, as we show later, significantly reduces the annotation burden. The method is of particular interest when building annotated databases for training deep-learning models, where the amount but also the diversity of the dataset are directly related to the performance and generalization ability.

4. EXPERIMENTAL VALIDATION

In order to validate the two main contributions of this work, we used the 2 recordings of 10 different patients over which the 3 muscles had sparse annotations. As in figure 6). We perform a comparison with other non-learning semi-automatic methods for muscle segmentation in 3D ultrasound images and evaluate the overlap of manual and the computed segmentation masks, as well as the difference in terms of volume estimation. Appendix A contains a more detailed description of the semi-automatic segmentation methods and evaluation metrics.

4.1 Qualitative results of label-propagation methods

We used the Slicer 3D open-source software implementation of the compared semi-automatic segmentation algorithms: "Fill Between Slices (FBS)", "Grow from seeds (GFS)" and "Watershed (WS)".

The results of these methods evidence the daily challenges that experts are faced with when provided manual annotations for ultrasound images. The lack of well-defined edges and the little contrast between regions of interest makes it difficult to define the segmentation mask borders for both experts and methods.

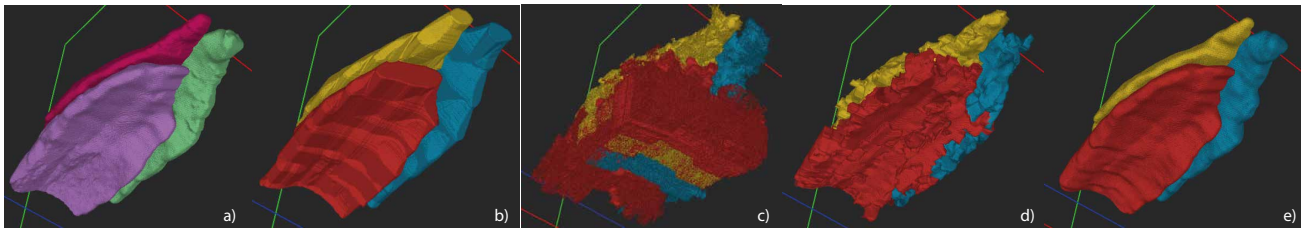

Figure 6. Qualitative results: a)Ground truth b)Fill between slices c)Grow from seeds d)Watershed e)Our method.

The qualitative visualization in Fig 6 shows that GFS and WS methods suffer from leakage. They have difficulties setting meaningful limits to the seed propagation due to the ambiguity of image contours on ultrasound images. FBS method ensures a smooth transition, but it demands parallel annotations that are not available in our case[‡]. Our dataset of 44 participants and 59 volumes has non-parallel partial masks, see Fig 4 b.

Our proposed ZOI method does not suffers from leakage and follows more closely the overall muscle shape. Nevertheless, it tends to over-smooth the results and some details on the borders may be missing.

4.2 Quantitative results of non-learning segmentation methods

In terms of segmentation performance, our method performs the best for the smaller muscles, GL and GM. The performance on the Soleus muscle is better with FBS followed by our method. We achieve 0.908 ± 0.04 of DICE and 0.877 ± 0.02 of mIoU over the amount of the muscles.

Our semi-automatic method is faster than slice by slice segmentation and FBS method, as we can see in table 1. It can be used as initialization for the remaining 34 acquisitions that have not been fully manually annotated. Initialization with the resultant volumes followed by expert corrections takes around 2 hours per muscle. This is less time-consuming compared with the 4 hours per muscle need on the reliable slice by slice method. This approaches brings a significant time gain.

[‡]We generate such parallel seeds by sampling from the the ground truth volumes an equivalent amount of slices to have a valid comparison (Fig 4 g)

Table 1. Quantitative Results for 10 participants

		DICE			IoU		
Algorithm	Time [min]	GL	GM	SOL	GL	GM	SOL
ZOI	90	**0.937**	**0.946**	0.841	**0.882**	**0.898**	0.791
FBS	96	0.909	0.924	**0.919**	0.835	0.859	**0.851**
GBS	44	0.818	0.755	0.763	0.699	0.615	0.620
WS	120	0.742	0.764	0.803	0.591	0.619	0.671

4.3 Volume estimation

In terms of volumetric error, our method presents smaller values, as shown in Table 2. we achieved around 4.17% of mean error computed over the transferred annotations x_1.

All of the non-learning semi-automatic segmentation methods on 3D US volumes provide quite good overlap with the volume computed from the ground-truth segmentation. For the Soleus muscle, Errors on the segmentation borders do not have a large impact on the estimated volume. Despite the leakage and muscles not being properly delimited for some methods, the volumetric error is reasonably low given the large region of overlap with the ground truth. In this sense, it is recommended that volume measurements are accompanied with additional measurements quantifying for instance the smoothness of the volume.

Table 2. Quantitative volumetric error for 10 participants

Algorithm	GL	GM	SOL
ZOI method	**2.61**	**4.98**	**4.93**
Fill between slices	10.00	10.70	13.03
Grow from seeds	24.02	10.61	15.53
Watershed	7.60	7.99	10.52

5. CONCLUSION AND FUTURE WORK

In this work, we propose a method consisting of simple steps to accelerate the manual segmentation of large volumetric US images. The expert is expected to provide sparse and partial annotations that are then propagated across a sweep of the US probe. A second part of the approach, acquires data with different frequencies and transfers labels from the simpler to the more complex to annotate acquisitions.

The simplicity of the zero-order interpolation approach makes it a good strategy for the application given the smooth changes of the muscle contours. The tested implementations of seed propagation methods suffer from leakage and would largely benefit from shape constraints, e.g. in the form of priors.

Given the importance of reducing annotation time, we will direct our method to fully annotate the 44 available US volumes, providing volumes with GM, GL, and Soleus muscles masks, on both low-frequency and high-frequency acquisitions.

These data will then serve for the training and validation of automatic deep-learning based segmentation methods.

ACKNOWLEDGMENTS

This work has been supported in part by the European Regional Development Fund, the Pays de la Loire region on the Connect Talent scheme (MILCOM Project) and Nantes Métropole (Convention 2017-10470). Special thanks to the teams working on the development of segmentation software as slicer, imfusion and stradwin.

REFERENCES

[1] Godi, C., Ambrosi, A., Nicastro, F., Previtali, S. C., Santarosa, C., Napolitano, S., Iadanza, A., Scarlato, M., Natali Sora, M. G., Tettamanti, A., et al., "Longitudinal mri quantification of muscle degeneration in duchenne muscular dystrophy," *Annals of clinical and translational neurology* **3**(8), 607–622 (2016).

[2] Le Guiner, C., Servais, L., Montus, M., Larcher, T., Fraysse, B., Moullec, S., Allais, M., François, V., Dutilleul, M., Malerba, A., et al., "Long-term microdystrophin gene therapy is effective in a canine model of duchenne muscular dystrophy," *Nature Communications* **8**, 16105 (2017).

[3] Pichiecchio, A., Alessandrino, F., Bortolotto, C., Cerica, A., Rosti, C., Raciti, M. V., Rossi, M., Berardinelli, A., Baranello, G., Bastianello, S., et al., "Muscle ultrasound elastography and mri in preschool children with duchenne muscular dystrophy," *Neuromuscular Disorders* **28**(6), 476–483 (2018).

[4] Pons, C., Borotikar, B., Garetier, M., Burdin, V., Ben Salem, D., Lempereur, M., and Brochard, S., "Quantifying skeletal muscle volume and shape in humans using mri: A systematic review of validity and reliability," *PloS one* **13**(11) (2018).

[5] Ihnatsenka, B. and Boezaart, A. P., "Ultrasound: Basic understanding and learning the language," *International journal of shoulder surgery* **4**(3), 55 (2010).

[6] Gonzalez Duque, V., AL CHANTI, D., Crouzier, M., Nordez, A., Lacourpaille, L., and Mateus, D., "Spatio-temporal Consistency and Negative Label Transfer for 3D freehand US Segmentation," (Oct. 2020).

[7] Yoshizumi, N., Saito, S., Koyama, D., Nakamura, K., Ohya, A., and Akiyama, I., "Multiple-frequency ultrasonic imaging by transmitting pulsed waves of two frequencies," *Journal of Medical Ultrasonics* **36**(2), 53–60 (2009).

[8] Che, C., Mathai, T. S., and Galeotti, J., "Ultrasound registration: A review," *Methods* **115**, 128–143 (2017).

[9] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al., "3d slicer as an image computing platform for the quantitative imaging network," *Magnetic resonance imaging* **30**(9), 1323–1341 (2012).

[10] Company, I., "Imfusion suite," (2020).

[11] of Cambridge, U., "Stradwin v6.02," (2020).

[12] Albu, A. B., Beugeling, T., and Laurendeau, D., "A morphology-based approach for interslice interpolation of anatomical slices from volumetric images," *IEEE Transactions on Biomedical Engineering* **55**(8), 2022–2038 (2008).

[13] Treece, G. M., Prager, R. W., Gee, A. H., and Berman, L., "Surface interpolation from sparse cross sections using region correspondence," *IEEE transactions on medical imaging* **19**(11), 1106–1114 (2000).

[14] Beare, R. and Lehmann, G., "The watershed transform in itk-discussion and new developments," *The Insight Journal* **1**, 1–24 (2006).

[15] Boykov, Y. Y. and Jolly, M.-P., "Interactive graph cuts for optimal boundary & region segmentation of objects in nd images," in [*Proceedings eighth IEEE international conference on computer vision. ICCV 2001*], **1**, 105–112, IEEE (2001).

[16] Crouzier, M., Lacourpaille, L., Nordez, A., Tucker, K., and Hug, F., "Neuromechanical coupling within the human triceps surae and its consequence on individual force-sharing strategies," *Journal of Experimental Biology* **221**(21) (2018).

[17] Zukic, D., Vicory, J., McCormick, M., Wisse, L., Gerig, G., Yushkevich, P., and Aylward, S., "Nd morphological contour interpolation," *Insight J* , 1–8 (2016).

[18] Lasso, A., "SlicerSegmentEditorExtraEffects:Watershed implementation," (2020).

[19] Soille, P., "Morphological image analysis: principles and applications," (2013).

APPENDIX A. DETAILS ON THE COMPARED METHODS

A.1 Non-learning methods for label propagation

- **Fill between slices**:
  It is a *iterative-morphological-Contour-Interpolator-method* proposed by Albu et al.[12] and implemented in 3DSlicer software.[17] It does not take into account the images of the volume, just the initial parallel spaced annotations.

  As input, it receives binary masks belonging to the same anatomical structure. Normally, structures evolve and regions split. It handles these topological changes. Each iteration interpolates between two shapes (A and B). Overlapping determine correspondences. A transition sequence of pixel morphological dilatation is generated.

  The m iterative conditional dilation of set B respect to A use the structuring element K. It is defined as:

  $$B \oplus_A^m K = (((B \oplus_A K) \oplus_A K)...)_{mtimes} \tag{1}$$

  Where $B \oplus_A K$ is defined as $((B \oplus K) \cap A.)$ This creates a smooth, gradual change of shape, without generating over-smooth.

  The implementation is multi-threaded, and processing one of the test inputs takes around 1 or 2 seconds on a quad-core processor. Normally, it is applied on the current transformer (CT) and MRI dataset composed of parallel slices. But it is also valid as an approach for our dataset of 3D ultrasound volumes.

- **Grow from seeds**:
  It is a *binary-interactive-automatic* segmentation method proposed by Boykov et al.[15] and implemented on 3DSlicer software.

  It takes into account the image $\mathbb{P}$, one seed of the object of interest ($\mathbb{o} \in \mathbb{O}$) and one seed of the background ($\mathbb{b} \in \mathbb{B}$) at least. The goal is to to find the optimal cut between background and foreground. A graph is defined. $\mathbb{G} =< \mathbb{E}, \mathbb{V} >$. Vertex are the pixels of the image and the seeds. In other worlds, vertex $v \in \mathbb{V}$ with $\mathbb{V} = \{p_i \in Image\} \cup \{o \in \mathbb{O}\} \cup \{b \in \mathbb{B}\}$.

  Each pixel has two links/edges at least, one to the background seed and other to the foreground seed ($\mathbb{e} \in \mathbb{E}$). The Value of edges is defined according to their type (neighborhood links or terminal links) with the table on.[15] The segmentation boundary is the minimum cost cut on the graph $\mathbb{G}$. It is computed exactly in polynomial time via algorithms for two terminal graph cuts.

  This interpolation method works well with homogeneous structures and contours well defined.

- **Watershed**:
  It is a *morphological-recursive* segmentation algorithm proposed by Beare et al.[14] and implemented on 3DSlicer software.[18] It uses markers, similar to grow from seeds, with the difference that smoothness of structures can be enforced, which can prevent leakage. In contrast, it is much slower.

  Gray-scale images are interpreted like "topographic maps", using intensities values as a new dimension. The brightest segments are the "borders", they delimit a sort of "catchment basins". The goal is to assign a different label to each basin, hopping they contain the object of interest.[19] On ultrasound, this focus property on the brightest regions is interesting because experts do the same. Bones' surfaces and Epymisiums [§] shine more.

  As input, it receives a grayscale image, seeds of the background, and seeds of, at least, one object of interest. Additionally, it receives a "percentage-value", it is an approximate percentage of pixels corresponding to the object of interest.

[§] Epymisium fibrous tissue envelope that surrounds the skeletal muscle.

Algorithm implementation first computes a bright spots detection over the grayscale image, saved as "feature-Image". It is the result of computing recursively a gradient magnitude filter after a Gaussian smoothness. Second, it creates a "label-image" with seeds values. Third, at least two priority queues are filled with neighboring pixels of seeds and the background. Priority level corresponds to the value on the "feature-image" (gradient magnitude). Fourth, neighbor pixels with the highest priority level are extracted from the priority queue, in the ratio defined with the "percentage-value". If extracted pixels are unlabeled, the value of the corresponding queue is assigned.

Steps three and four are repeated until priority queues are empty. This floods successively the grey value relief, and basins emerge along the edges. Over medical images this algorithm normally over-segment. Either the image must be pre-processed or the regions must be merged on the basis of a similarity criterion afterward.

A.2 Metrics for 3D volumes

Prediction errors for 3D volumes can be evaluated in terms of the contour, the area, and the content. Contour metrics use the surface correlation. Area metrics use values of true positives, true negatives, false positive and false negatives pixels (TP, TN, FP, FN) to define a percentage of accuracy. Content metrics calculate the existence of inside holes and boundary holes in the segmented region.

- **Dice and IoU**:
  These metrics evaluate the correlation between the ground truth and the prediction. It use the amount of true positive values(TP), false positive values(FP) and false negative values(FN).

$$dice = \frac{2 * TP}{2TP + FP + FN} \tag{2}$$

$$IoU = \frac{dice}{2 - dice} \tag{3}$$

- **Average Surface Distance and Hausdorff distance**:
  These metrics estimates the error between the surfaces of the ground truth and the prediction (S and S'). If $p \in S$, the distance to S' is given by the minimum of the Euclidean norm: $d(p, S') = min_{p' \in S'} ||p - p'||_2$. Calculating these values for all the points, with n_S equal to the number of pixels on S surface, the average surface distance is given by equation 4. Hausdorff distance is the largest difference between the surface distances, expressed in equation 5. To obtain values in mm, we multiply metrics by the spacing of the volume.

$$ASD = (\frac{1}{n_S + n_{S'}}) * (\sum_{p=1}^{n_S} d(p, S') + \sum_{p'=1}^{n_{S'}} d(p', S)) \tag{4}$$

$$HD = max[d(S, S'), d(S', S)] \tag{5}$$