Fetal Body Parts Segmentation Using Volumetric MRI Reconstructions

Pedro Pablo Alarcón-Gil[1(✉)], Felicia Alfano[1], Alena Uus[2], María Jesús Ledesma-Carbayo[1,3], and Lucilio Cordero-Grande[1,2,3]

[1] Biomedical Image Technologies, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain
pedro.alarcon.gil@upm.es

[2] Center for the Developing Brain, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK

[3] Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, Spain

Abstract. Fetal body parts segmentation can be useful to detect abnormalities and assess fetal growth from magnetic resonance imaging (MRI). In this work, 3D fetal head, trunk and limbs body parts segmentation is leveraged for the first time by volumetric reconstructions of the whole fetal anatomy coupled to deep learning techniques for volumetric segmentation. Due to the time consuming manual segmentation required for training in this volumetric multi-label setting, sparse annotations are used, marking slices with a 6 mm separation and alternating the slicing axis between cases. These manual segmentations are used to train and compare different models using both the original MRI data and the data after volumetric reconstruction in a dataset of 45 cases. The setup consisting on 3D volumetric reconstructions and a 3D U-net based learning model results in optimal segmentation metrics, with Dice scores higher than 0.9 for all the considered structures and 0.973 for the fetal body (0.979 when highly motion corrupted datasets are discarded). In comparison, the performance of the setups that use the original MRI data exhibit a pronounced decline in segmentation scores, highlighting the importance of robust reconstruction techniques for automatic fetal growth characterization. Finally, we conduct a Bland-Altman analysis studying the reliability of our proposed 3D reconstruction and segmentation pipeline for automatic estimation of fetal body part weights.

Keywords: Fetal MRI · Body segmentation · Volumetric reconstruction · Deep learning · Fetal growth biomarkers

1 Introduction

Magnetic resonance imaging (MRI) is a non invasive diagnostic imaging modality recommended when fetal abnormalities are suspected by ultrasound (US) [1]. Fetal MRI has become a well-accepted adjunct due to advantages over US in

D. Link-Sourani et al. (Eds.): PIPPI 2024, LNCS 14747, pp. 130–140, 2025.
https://doi.org/10.1007/978-3-031-73260-7_12

terms of larger field of view (FOV), lower operator skills dependence, better soft tissue contrast or absence of shadowing [2]. In typical fetal MRI acquisitions the data is collected slice by slice in stacks oriented along one or more axes. After that, the acquired slices could be used to construct a volumetric image, but significant interslice inconsistencies are typically present associated with fetal motion and mother breathing during the examination.

Slice to volume reconstruction (SVR) methods have been proposed for volumetric representation of the imaged anatomy with reduced inconsistencies. Most methods have focused on rigid motion of the fetal brain, but recent proposals have extended this framework to elastic motion models for fetal trunk and placenta [3,4] or whole uterus [5] reconstruction. Fetal trunk reconstruction allows the 3D characterization of various structures including the lungs, thymus, cardiac vessels, liver or kidneys, and enables the generation of fetal growth curves for specific organs, as well as atlas construction [6]. However, to the best of our knowledge, no attempts have been made to leverage volumetric deformable SVR for consistent whole fetal body segmentation and its division into head, trunk and limbs as a step for more comprehensive analyses of fetal growth.

Fetal body segmentation is of great interest in various applications. Automatic segmentation procedures could enable more reliable estimates of fetal weight (EFW) and facilitate the segmentation of fetal body structures. Recent investigations in fetal body segmentation have suggested MRI as an alternative to US for EFW [7], leading to fetal growth charts tailored specifically for MRI-based EFW assessments [8]. In addition, if fetal body segmentations are available, correlations between fetal organ volumes and specific diseases can not only be established considering absolute volumes of the structures of interest, but also their relative volumes in relation to total body volume [9,10].

Several automatic segmentation approaches have been introduced to address fetal body segmentation [7,8,11–15]. Preceding the deep learning (DL) emergence, [11] employed a graph-based segmentation method with modest performance, yielding Dice scores lower than 0.7. Within the DL paradigm, performance has increased significantly. However, data and techniques used for training the models become crucial in this context. Obtaining a reliable ground truth (GT) is a time consuming task, so strategies based on sparse and semi-automatic annotation have been proposed. Sparse annotation is used in [12] with a 5 mm slice separation and in [15] by segmenting small patches around a randomly selected slice inside previously delineated fetal body limits. Semi-automatic annotation is used in [8] where a network is trained with a few cases to obtain a first outline, and posterior operator review is performed to generate the final GT. The data used for training the models should support the widest generalizability for further application to novel datasets. For this reason, [7] proposed 2D slices from all scan axes for training and a majority voting mechanism for inference, contrasting the conventional practice of using single axis data [13,14].

Largely due to limitations in volumetric consistency of image data used for segmentation, previous approaches only confront the whole body segmentation problem. In this paper, we not only address the binary fetal body segmenta-

tion, but also the multi-label segmentation into its main parts; namely, head, trunk, and limbs, which is leveraged by 3D reconstruction and segmentation. Our method achieves better performance in whole fetal body segmentation using full uterus SVR, thus suggesting a tool for improved EFW reliability. In addition, our multi-part segmentations could be used to resolve the fetal pose, so that fetal scans can all be reformatted to a fetal body oriented coordinate system, thereby facilitating assessments. Moreover, the segmented components allow fetal growth characterization of distinct body parts, provide biomarkers and renderings useful to detect fetal growth anomalies, and could be used to constrain and enrich subsequent analysis steps.

2 Materials and Methods

This study includes 45 fetal cases consented as participants in the iFIND project [16] with gestational ages (GAs) distribution ranging from 20 to 36 weeks. Images were acquired on a Philips Ingenia 1.5 T with a 24-channel receive coil using a single-shot fast spin echo sequence. 5 stacks were acquired per subject, including one axial, two sagittal and two coronal repeats, each with appropriate number of slices for tight uterus coverage. Data was collected with an in-plane resolution of 1.25 mm isotropic, 2.5 mm slice thickness and 1.25 mm slice separation. Sensitivity encoding acceleration was set to 2 with half scan 0.575 and echo time of 80 ms. Total acquisition time was on average $11'20''$.

We constructed two GTs for fetal body segmentation: one based on Raw-MRI (the original data) and another on SVR-MRI (the data resulting from volumetric reconstructions). For volumetric image generation, we used robust generative diffeomorphic SVR [5], which, up to our knowledge, is the only solution providing consistent reconstructions of whole fetuses, including head, trunk and limbs, and their uterine context. To reduce manual marking burden per case, we performed sparse annotations of slices separated by 6 mm. We alternated the axis onto which our manual segmentations are subsampled (be it transversal, coronal or sagittal, all referred to the acquisition geometry) for the different cases. One stack is annotated through the selected axis in Raw-MRI, and the same marking axis is used for SVR-MRI (see Fig. 1). Using a in-house tool, annotations were reviewed with interpolated volumetric surfaces visualization.

The generated GTs were used to train a convolutional neural network (CNN). We adapted the original nnUNet framework [17] as the workhorse for our tests, given that state of the art (SOTA) results have been recently confirmed in several 2D and 3D benchmarks [18] and it has been successfully applied to whole body fetal segmentation before [8]. Both 2D and 3D nnUNet models have been tested as no clear a priori choice appeared possible due to unpredictable performance impact of motion-induced volumetric inconsistencies. The slices marked in each volume were directly used for 2D model training. For 3D model training the nnUNet implementation was extended to disable unmarked slices backpropagation. Of note, self-configuration of nnUNet provided similar training times for 2D and 3D models. Default nnUNet parameters were used in all tests, except

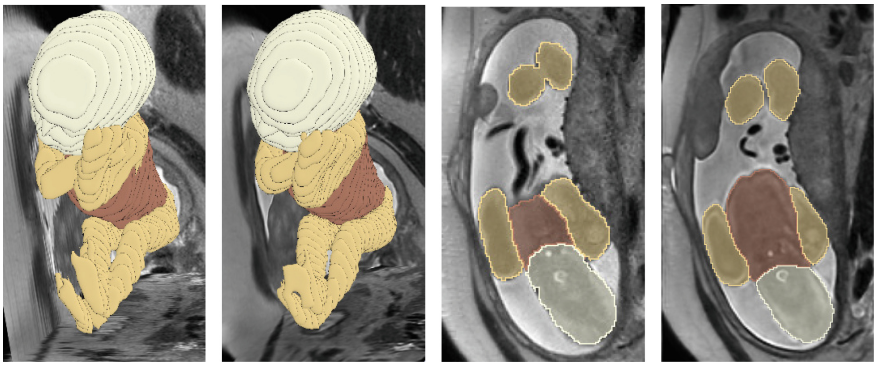

Fig. 1. Illustration of 6 mm separation in slice annotations. 3D view in (a) Raw-MRI and (b) SVR-MRI. 2D view in (c) Raw-MRI and (d) SVR-MRI.

for the number of epochs which was reduced to 100 to expedite comparisons. All models were trained using a NVIDIA RTX A6000.

A comparison between 2D and 3D model training using both Raw-MRI and SVR-MRI data was conducted based on our marked GT using 5-fold cross validation. Given that no hyperparameter tuning is performed on validation, presented results correspond to 5 training runs averaged in 5 independent test sets. The same splits, each comprising 36 cases for training and remaining 9 cases for testing, were used to train and test all the models. Inference and postprocessing were performed as shown in Fig. 2. Inference depends on the network dimension and input data nature. For Raw-MRI –Fig. 2(a) and (b)–, network predictions are obtained over the 5 views and majority voting is used to obtain a unique segmentation. For SVR-MRI –Fig. 2(c) and (d)–, predictions are directly obtained from the input volume. In 2D schemes –Fig. 2(a) and (d)–, 2D slicing is applied before inference and results are stacked into a 3D segmentation. Identical postprocessing is applied to final 3D results in all configurations, consisting in largest connected component extraction for the fetal body and removal of internal holes for each of the three body parts regions.

Validation compares previous configurations using Dice coefficients (DC), normalised surface distances (NSD) with 2 mm tolerance, and mean absolute surface distances (MASD). Manually marked slices are directly used for the regional DC measure, while for NSD and MASD, we consider the annotation anisotropy when performing distance computations. For Raw-MRI, we compare models with the aforementioned majority voting and also considering inference in the annotated repeat only (Raw-MRI-1-view), to deduce errors due to inter-stack motion. Statistical significance of the best ranked configuration is tested using a Bonferroni multiple comparison corrected Kruskal-Wallis test. We also report results when excluding cases with severe motion artifacts (6 out of 45 cases), referred to as high-quality (HQ) scans. A visual comparison of the three

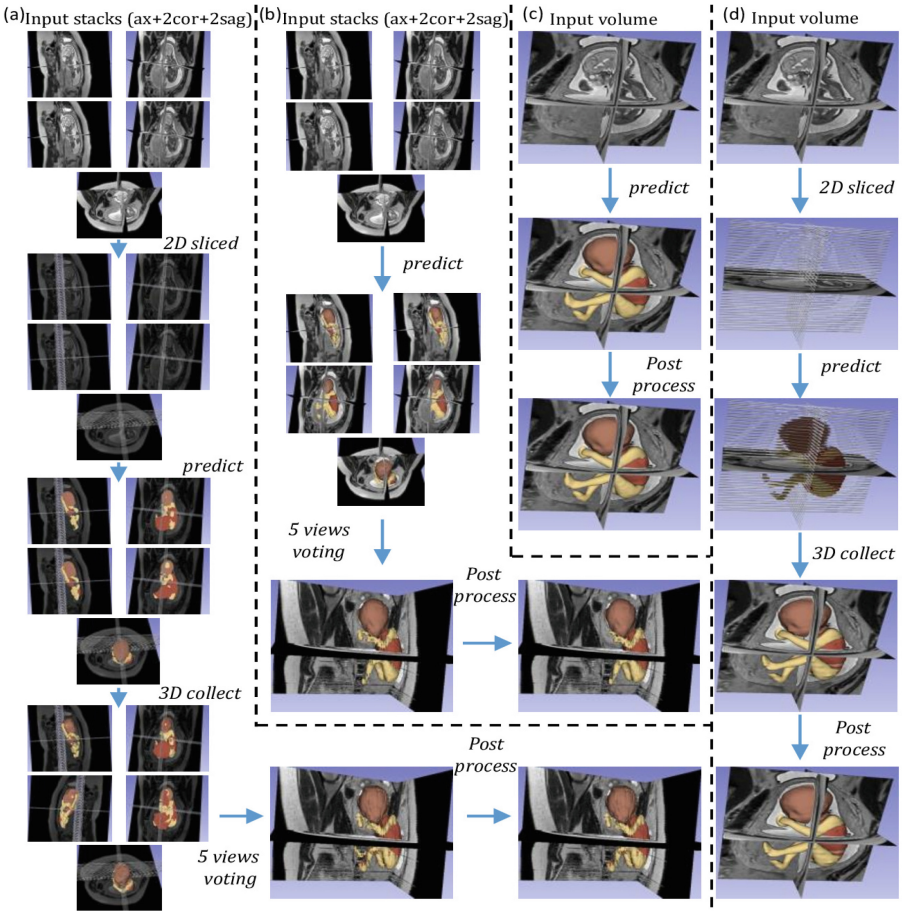

Fig. 2. Inference and postprocessing flowchart for (a) 2D Raw-MRI, (b) 3D Raw-MRI, (c) 3D SVR-MRI and (d) 2D SVR-MRI.

best configurations is presented using the cases corresponding to the 25 and 75 percentiles of the averaged body DC (respectively showcasing poor and good performances). Bland-Altman (BA) EFW analysis is used to compare regional volumes from manual and automatic segmentations in HQ scans. The EFW biomarker is determined by application of the Baker's formula [19] to fetal body and body parts volumes computed as the segmentation voxel counts with annotation anisotropy correction.

3 Results

Models using SVR-MRI get better DC (Table 1) and NSD (Table 2) than those using Raw-MRI, with the exception of similar performance of 2D SVR-MRI compared to 3D Raw-MRI-1-view for the fetal body parts. In addition, 3D SVR-MRI

proves to be the best configuration for all reported metrics, including also MASD (Table 3), and achieves statistical significance ($p < 0.05$) in all cases except when compared to 2D SVR-MRI for fetal body segmentations. Misclassifications of main fetal body parts in 2D Raw-MRI-1-view and 2D SVR-MRI are suggested by the relatively larger differences in MASD between these configurations and their 3D counterparts when referenced to the fetal body results. Limbs performance is worse than trunk performance for MASD and NSD in Raw-MRI, but this trend is reversed in SVR-MRI due to improved limbs consistency after reconstruction. Classic problems for fetal volumetric imaging are also apparent from these results. Fetal motion is behind the worse DC and NSD for Raw-MRI over Raw-MRI-1-view, showing the mismatch between other stacks and the one used to generate the GT. Interslice inconsistencies in Raw-MRI are evidenced by 2D Raw-MRI having better performance and reduced variance compared to 2D Raw-MRI-1-view in MASD, due to discontinuous fetal body parts segmentations in the latter, which often induce heavy under-segmentation after postprocessing. Finally, in the last row of each table, we present 3D SVR-MRI results when discarding highly artifacted datasets, showing the expected performance of our proposed reconstruction and segmentation pipeline if HQ input data is available.

Table 1. Mean ± std(p-value) for trunk, head, limbs and whole body Dice coefficients. p-value refers to Kruskal-Wallis test with respect to 3D SVR-MRI.

Dice	Head	Trunk	Limbs	Body
2D-Raw	$0.900 \pm 0.022 (< 10^{-9})$	$0.843 \pm 0.033 (< 10^{-9})$	$0.561 \pm 0.051 (< 10^{-9})$	$0.858 \pm 0.024 (< 10^{-9})$
2D-Raw-1view	$0.920 \pm 0.022 (< 10^{-9})$	$0.863 \pm 0.034 (< 10^{-9})$	$0.718 \pm 0.045 (< 10^{-9})$	$0.909 \pm 0.024 (< 10^{-9})$
3D-Raw	$0.901 \pm 0.020 (< 10^{-9})$	$0.847 \pm 0.042 (< 10^{-9})$	$0.606 \pm 0.050 (< 10^{-9})$	$0.862 \pm 0.027 (< 10^{-9})$
3D-Raw-1view	$0.938 \pm 0.009 (< 10^{-8})$	$0.903 \pm 0.010 (< 10^{-6})$	$0.783 \pm 0.024 (< 10^{-6})$	$0.930 \pm 0.009 (< 10^{-8})$
2D-SVR	$0.944 \pm 0.011 (< 10^{-5})$	$0.904 \pm 0.013 (< 10^{-5})$	$0.816 \pm 0.024 (0.001)$	*0.958 ± 0.006(0.160)*
3D-SVR	**0.975 ± 0.007**	**0.951 ± 0.010**	**0.907 ± 0.017**	**0.973 ± 0.006**
3D-SVR-HQ	0.982 ± 0.002	0.958 ± 0.005	0.921 ± 0.009	0.979 ± 0.002

Table 2. Mean ± std(p-value) for trunk, head, limbs and whole body NSD. p-value refers to Kruskal-Wallis test with respect to 3D SVR-MRI.

NSD	Head	Trunk	Limbs	Body
2D-Raw	$0.774 \pm 0.047 (< 10^{-9})$	$0.639 \pm 0.040 (< 10^{-9})$	$0.562 \pm 0.038 (< 10^{-9})$	$0.697 \pm 0.034 (< 10^{-9})$
2D-Raw-1view	$0.819 \pm 0.046 (< 10^{-8})$	$0.693 \pm 0.038 (< 10^{-9})$	$0.709 \pm 0.045 (< 10^{-9})$	$0.829 \pm 0.037 (< 10^{-8})$
3D-Raw	$0.750 \pm 0.051 (< 10^{-9})$	$0.641 \pm 0.047 (< 10^{-9})$	$0.579 \pm 0.037 (< 10^{-9})$	$0.696 \pm 0.038 (< 10^{-9})$
3D-Raw-1view	$0.859 \pm 0.026 (< 10^{-7})$	$0.749 \pm 0.029 (< 10^{-5})$	$0.776 \pm 0.026 (< 10^{-5})$	$0.856 \pm 0.022 (< 10^{-7})$
2D-SVR	$0.862 \pm 0.029 (< 10^{-6})$	$0.742 \pm 0.027 (< 10^{-6})$	$0.808 \pm 0.026 (0.003)$	*0.923 ± 0.016(0.179)*
3D-SVR	**0.957 ± 0.021**	**0.882 ± 0.023**	**0.913 ± 0.018**	**0.963 ± 0.012**
3D-SVR-HQ	0.974 ± 0.008	0.895 ± 0.019	0.925 ± 0.013	0.972 ± 0.008

Visual comparisons are shown in Fig. 3 for poor –Fig. 3(a)– and good –Fig. 3(b)– performance examples (corresponding to 25 and 75 segmentation quality percentiles) using 3D Raw-MRI, 2D SVR-MRI and 3D SVR-MRI. Fetal

Table 3. Mean ± std(p-value) for trunk, head, limbs and whole body MASD(mm). p-value refers to Kruskal-Wallis test with respect to 3D SVR-MRI.

MASD(mm)	Head	Trunk	Limbs	Body
2D-Raw	$1.942 \pm 0.777(< 10^{-9})$	$2.709 \pm 0.568(< 10^{-9})$	$3.598 \pm 0.489(< 10^{-9})$	$2.173 \pm 0.318(< 10^{-9})$
2D-Raw-1view	$2.029 \pm 0.809(< 10^{-9})$	$3.066 \pm 0.766(< 10^{-9})$	$3.878 \pm 1.337(< 10^{-9})$	$2.114 \pm 0.871(< 10^{-9})$
3D-Raw	$1.647 \pm 0.398(< 10^{-9})$	$2.717 \pm 0.834(< 10^{-9})$	$3.587 \pm 0.537(< 10^{-9})$	$2.248 \pm 0.371(< 10^{-9})$
3D-Raw-1view	$0.887 \pm 0.180(< 10^{-6})$	$1.683 \pm 0.253(< 10^{-3})$	$1.798 \pm 0.305(< 10^{-4})$	$1.031 \pm 0.185(< 10^{-6})$
2D-SVR	$1.462 \pm 0.494(< 10^{-9})$	$2.409 \pm 0.365(< 10^{-9})$	$1.797 \pm 0.392(< 10^{-4})$	*0.618 ± 0.188(0.162)*
3D-SVR	**0.312 ± 0.143**	**0.978 ± 0.344**	**0.725 ± 0.223**	**0.314 ± 0.105**
3D-SVR-HQ	0.204 ± 0.051	0.813 ± 0.225	0.566 ± 0.101	0.235 ± 0.054

motion issues can be observed in 3D Raw-MRI, which tends to produce rough, noisy or incomplete limbs segmentations –see the area enclosed by the red circle in Fig. 3(b)–. 2D SVR-MRI shows artifacted head segmentation for the poor performance example and head and trunk swaps for the good performance example –see the area pointed by the red arrow in Fig. 3(b)–. In contrast, 3D SVR-MRI results provide smoother and more reliable results for the pinpointed examples.

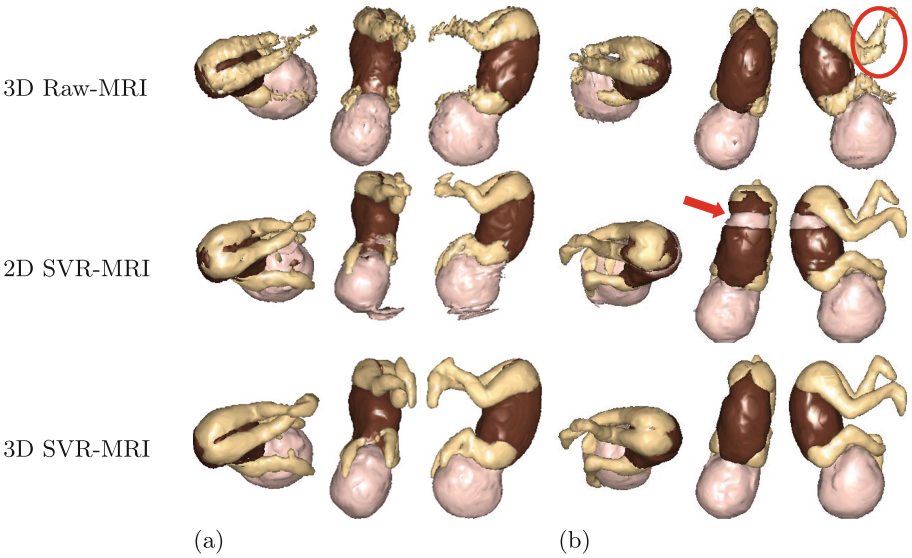

Fig. 3. Visual comparison of 3D Raw-MRI, 2D SVR-MRI and 3D SVR-MRI segmentations for (a) poor and (b) good performance examples in the main scan axes.

BA analysis of EFW of the main body parts using manual and 3D SVR-MRI-HQ segmentations (Fig. 4) demonstrates a strong agreement, with negligible biases for head and trunk. However, smaller automatic than manual EFW is observed for the limbs, which has a certain impact on whole body estimates. Variance and corresponding Limits of Agreement (LoA) of EFW are larger for

trunk and limbs than for body and head, with similar LoA observed for trunk and limbs suggesting possible classification swaps between these two regions. Biases and LoA for the different regions are reported in the caption of Fig. 4. We observe that agreements of EFW between manual and automatic segmentations do not present notable differences across GAs.

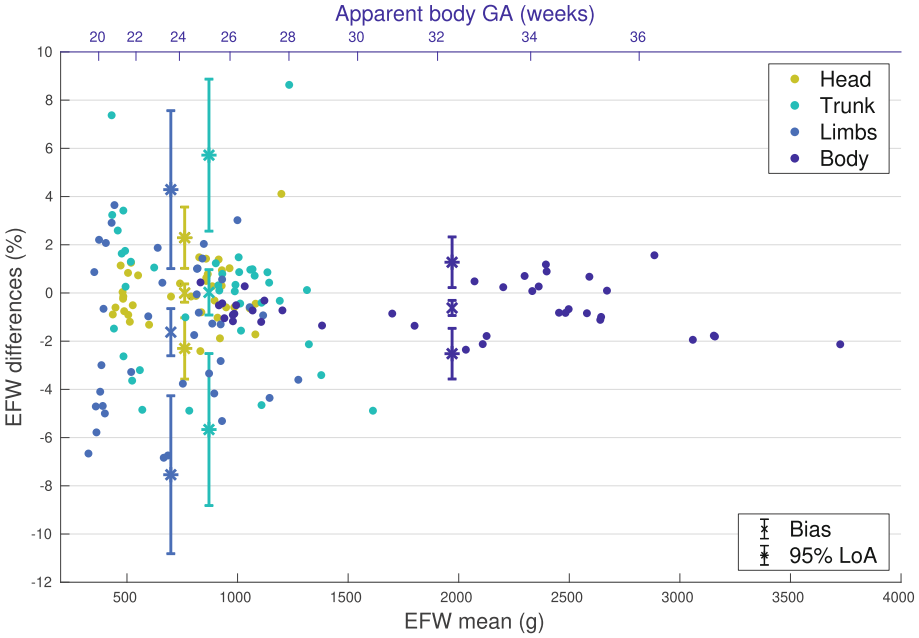

Fig. 4. BA analysis of EFW using 3D SVR-MRI-HQ and manual annotations for the head, trunk, limbs and whole body. Horizontal bottom axis displays the mean of both EFW. Horizontal top axis represents apparent body GA obtained by a logarithmic regressor constructed using clinical GA information. This is included to illustrate association of whole body EFW to corresponding stages of gestation. Vertical axis shows relative automatic vs manual EFW differences. Cross error bars indicate biases and corresponding confidence intervals. Asterisk error bars indicate 95% LoA and corresponding confidence intervals. biases $\pm$ 95%LoA: $-0.004 \pm 2.297\%$ (head), $0.026 \pm 5.692\%$ (trunk), $-1.626 \pm 5.914\%$ (limbs) and $-0.622 \pm 1.897\%$ (body).

4 Discussion

Our results show that 3D SVR-MRI is the optimal fetal body and body parts segmentation pipeline. Its superiority is attributed to SVR-MRI increased image quality compared to Raw-MRI and to 3D architectures being more appropriate than 2D counterparts for fetal body part segmentation. Interestingly, 2D SVR-MRI outperforms 3D Raw-MRI in all metrics for whole fetal body segmentation,

suggesting that reconstruction volumetric consistency plays a more relevant role than network dimensionality for this problem, with issues derived from low quality slices in the acquisition illustrated by results in the area enclosed in the red circle in Fig. 3(b). However, the superiority of 2D SVR-MRI compared to 3D Raw-MRI is not so clear for the main fetal body parts, as 2D models often have limited contextual information to determine body parts from independent slices they see, as shown in the misclassification pointed by the arrow in Fig. 3(b).

Our 3D SVR-MRI model achieves fetal body segmentation SOTA performance (DC over 0.97) using fewer training epochs (100 versus 1000) and training cases (36 versus 66) than previous leading works [8]. However, comparisons are difficult as GTs are constructed manually in our case and semi-automatically in [8], and the balance of motion artifacted cases in different cohorts can have a decisive impact in reported performance. Nevertheless, the high DC reported in SOTA fetal body segmentation may suggest limited potential for improvement, at least in terms of accuracy. For this reason, we have investigated an extended segmentation problem aimed at differentiating the head, trunk and limbs.

SVR-MRI effectively leverages 3D information from acquired data, with volumetric consistency of reconstructions translating into more consistent segmentations. Anyhow, limbs segmentation exhibits inferior DC scores than trunk and head segmentation, which can be attributed to fetal motion and substantial variability in limb poses. Also, this has an impact for volumetry, as the limbs present the highest combined bias and variance in the BA analysis.

Future extensions of our work could include using the fetal body main parts segmentations to retrospectively reformat SVR-MRI reconstructions to fetal body coordinates as an initial step for further analysis of fetal growth in anatomically matched coordinates. In the fetal spatial coordinates, it might be easier to address limbs differentiation, although this may face technical challenges due to the limitations of CNNs in dealing with laterality and verticality [20]. Another future direction may involve the localized segmentation of additional fetal body structures such as the organs inside the trunk.

5 Conclusions

We have presented a multi-class fetal body parts segmenter based on volumetric reconstructions and DL. The use of SVR-MRI induces notable improvements in the segmentation of the fetal body, especially when considering also the body parts segmentation problem. In terms of model dimension, 3D models have been shown to generally surpass 2D models in our experiments. Finally, our 3D-SVR-MRI-HQ automatic pipeline shows an excellent agreement with manual EFW biomarker extraction, and can be used to enrich image-based characterization of fetal growth.

Acknowledgments. This work was supported in part by MCIN, Spain, under the Beatriz Galindo Program (BGP18/00178) and the STEP-AMD project (TED2021-1319518-I00); in part by MCIN/AEI/10.13039/5011000110 33/FEDER, EU, under

Projects PID2021-129022OA-I00 and PID2022-141493OB-I00; and in part by the Madrid Government, Spain, under the Multiannual Agreement with Universidad Politécnica de Madrid in the line support for Research and Development Projects for Beatriz Galindo Researchers, in the context of V PRICIT. The authors gratefully acknowledge the Universidad Politécnica de Madrid for providing computing resources on Magerit Supercomputer.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Papaioannou, G., Caro-Domínguez, P., Klein, W.M., Garel, C., Cassart, M.: Indications for magnetic resonance imaging of the fetal body (extra-central nervous system): recommendations from the European Society of Paediatric Radiology Fetal Task Force. Pediatr. Radiol. **53**(2), 297–312 (2023). https://doi.org/10.1007/s00247-022-05495-4
2. Wilson, L., Whitby, E.H.: The value of fetal magnetic resonance imaging in diagnosis of congenital anomalies of the fetal body: a systematic review and meta-analysis. BMC Med. Imag. **24**(1) (2024). https://doi.org/10.1186/s12880-024-01286-5
3. Uus, A., et al.: Deformable slice-to-volume registration for motion correction of fetal body and placenta MRI. IEEE Trans. Med. Imaging **39**(9), 2750–2759 (2020). https://doi.org/10.1109/tmi.2020.2974844
4. Mufti, N., et al.: Use of super resolution reconstruction MRI for surgical planning in Placenta accreta spectrum disorder: case series. Placenta **142**, 36–45 (2023). https://doi.org/10.1016/j.placenta.2023.08.066
5. Cordero-Grande, L., et al.: Fetal MRI by robust deep generative prior reconstruction and diffeomorphic registration. IEEE Trans. Med. Imaging **42**(3), 810–822 (2023). https://doi.org/10.1109/tmi.2022.3217725
6. Uus, A.U., et al.: Automated body organ segmentation, volumetry and population-averaged atlas for 3D motion-corrected T2-weighted fetal body MRI. Sci. Rep. **14**(1) (2024). https://doi.org/10.1038/s41598-024-57087-x
7. Ryd, D., Nilsson, A., Heiberg, E., Hedström, E.: Automatic segmentation of the fetus in 3D magnetic resonance images using deep learning: accurate and fast fetal volume quantification for clinical use. Pediatr. Cardiol. **44**(6), 1311–1318 (2023). https://doi.org/10.1007/s00246-022-03038-0
8. Specktor-Fadida, B., et al.: Deep learning-based segmentation of whole-body fetal MRI and fetal weight estimation: assessing performance, repeatability, and reproducibility. Eur. Radiol. **34**(3), 2072–2083 (2024). https://doi.org/10.1007/s00330-023-10038-y
9. Rabinowich, A., et al.: Fetal MRI-based body and adiposity quantification for small for gestational age perinatal risk stratification. J. Magn. Reson. Imaging **60**(2), 767–774 (2024). https://doi.org/10.1002/jmri.29141
10. Hall, M., et al.: Adrenal volumes in fetuses delivering prior to 32.weeks' gestation: an MRI pilot study. Acta Obstetricia et Gynecologica Scandinavica **103**(3), 512–521 (2024). https://doi.org/10.1111/aogs.14733
11. Zhang, T., et al.: Graph-based whole body segmentation in fetal MR images. In: MICCAI Workshop PIPPI (2016). https://pippiworkshop.github.io/pippi2016/pdf/PIPPI2016_04_Zhang.pdf

12. Kulseng, C.P.S., Hillestad, V., Eskild, A., Gjesdal, K.I.: Automatic placental and fetal volume estimation by a convolutional neural network. Placenta **134**, 23–29 (2023). https://doi.org/10.1016/j.placenta.2023.02.009
13. Dudovitch, G., Link-Sourani, D., Ben Sira, L., Miller, E., Ben Bashat, D., Joskowicz, L.: Deep learning automatic fetal structures segmentation in MRI scans with few annotated datasets. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12266, pp. 365–374. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59725-2_35
14. Lo, J., et al.: Cross attention squeeze excitation network (CASE-Net) for whole body fetal MRI segmentation. Sensors **21**(13) (2021). https://doi.org/10.3390/s21134490
15. Fadida, B.S., Sourani, D.L., Sira, L.B., Miller, E., Bashat, D.B., Joskowicz, L.: Partial annotations for the segmentation of large structures with low annotation cost. In: Zamzmi, G., Antani, S., Bagci, U., Linguraru, M.G., Rajaraman, S., Xue, Z. (eds.) MILLanD 2022. LNCS, vol. 13559, pp. 13–22. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16760-7_2
16. iFIND: Intelligent fetal imaging and diagnosis. http://www.ifindproject.com
17. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
18. Isensee, F., et al.: nnU-Net revisited: a call for rigorous validation in 3D medical image segmentation (2024). https://doi.org/10.48550/arXiv.2404.09556
19. Baker, P., et al.: Fetal weight estimation by echo-planar magnetic resonance imaging. Lancet **343**(8898), 644–645 (1994). https://doi.org/10.1016/s0140-6736(94)92638-7
20. Liu, R., et al.: An intriguing failing of convolutional neural networks and the CoordConv solution, July 2018. https://doi.org/10.48550/arXiv.1807.03247