Multi-modal deep learning for fetal corpus callosum segmentation and characterization in MRI

Marina Di Stefano[1,2] [0009−0001−0711−6263], Denis Peruzzo[1] [0000−0002−9480−379X], Florian Montano[3] [0009−0004−7249−3796], Roxane Licandro[4,5,6,7,8,9] [0000−0001−9066−4473], Alberto De Luca[2] [0000−0002−2553−7299], Sonja MC De Zwarte[10] [0000−0001−9015−3550], Alexander Leemans[2] [0000−0002−9306−6126], and Tommaso Ciceri[1] [0000−0001−5566−8110]

tommaso.ciceri@lanostrafamiglia.it

[1] Neuroimaging Unit, Scientific Institute IRCCS E. Medea, Bosisio Parini, Italy
[2] Image Sciences Institute, Division Imaging and Oncology, University Medical Center Utrecht, Utrecht, the Netherlands
[3] Diagnostic Imaging and Neuroradiology Unit, Scientific Institute IRCCS E. Medea, Bosisio Parini, Italy
[4] Computational Imaging Research Lab, ELIA Group, Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna (MedUni Wien), Austria
[5] Comprehensive Center for Artificial Intelligence in Medicine, MedUni Wien, Austria
[6] Christian Doppler Laboratory for Mathematical Modelling and Simulation of Next-Generation Medical Ultrasound Devices, Austria
[7] Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[8] Department of Early Life Imaging, School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK
[9] Smart Imaging Lab, University Hospital Erlangen, Erlangen, Germany
[10] Department of Developmental Psychology, Utrecht University, Utrecht, the Netherlands

Abstract. The corpus callosum (CC) plays a fundamental role in inter-hemispheric integration, supporting motor coordination, language, and sensory processing. The ability to accurately assess the morphology of CC prenatally would allow the establishment of normative neurodevelopmental trajectories. Automated fetal brain tissue segmentations typically rely on T2-weighted (T2w) MR images, however limited contrast between the developing CC and adjacent structures hampers precise delineation. We proposed a multi-modal deep learning model that combines anatomical (T2w) and microstructural (fractional anisotropy, FA) information to address this limitation. Leveraging 192 fetal brain scans of the developing Human Connectome Project from 21 to 38 gestational weeks, we generated high-fidelity ground-truth annotations using an initialization pipeline that combines morphological filtering of FA and color FA images. We then systematically evaluated three segmentation models based on T2w, FA, and combined ($T2w + FA$) in-

puts. The dual-channel ($T2w + FA$) model significantly outperformed state-of-the-art methods, achieving a Dice score of 0.74 ± 0.07, volume similarity of 0.92 ± 0.06, and perfect topological consistency (Euler Difference = 0). Furthermore, performance remained quite stable under a simulated reduced-acquisition protocol with fewer diffusion-weighted directions, suggesting potential applicability to lower-resolution clinical diffusion MRI, though further validation on external sites and routine clinical-quality acquisitions is needed. Morphometric analysis revealed robust gestational age-dependent growth trajectories for most of the extracted parameters, while FA remained stable throughout gestation (Spearman $\rho = -0.02$), suggesting that the microstructural properties underlying callosal anisotropy change little over this developmental period. This multi-modal model offers a robust tool for prenatal CC assessment, enabling quantitative characterization of typical development and facilitating the identification of derailments from the expected developmental trajectory. Code and trained models are publicly available at https://github.com/tommaso-ciceri/Multi-modal-deep-learning-for-fetal-MRI-corpus-callosum-segmentation.

Keywords: Fetal brain · Corpus callosum · Segmentation · DTI.

1 Introduction

The corpus callosum (CC), the principal interhemispheric pathway in the brain, plays a central role in neurodevelopment, being implicated in numerous higher-order functions. Abnormalities of the CC are among the most common malformations of the central nervous system, with an estimated prevalence of 0.3–0.7% in fetuses and up to 2–3% in developmentally disabled patients [5, 16]. Although prenatal ultrasound is the primary screening modality for CC anomalies, its high false-positive rate has established fetal MRI as the confirmatory reference standard [5].

Due to the CC's importance in neurodevelopment, there is growing interest in automated methods for its characterization in prenatal imaging, with segmentation serving as a key step for isolating this structure and enabling its quantitative morphological analysis [11, 27]. These approaches could reduce the time of manual annotations while enabling rapid, reproducible extraction of quantitative biomarkers (e.g. thickness, area). Currently, most fetal brain segmentation methods are based on the gold-standard U-Net–type architectures and are trained on manually labeled data [3]. CC annotations usually rely on T2-weighted (T2w) MR scans, which are suboptimal in the fetal brain: the CC is thin and has low contrast with the surrounding parenchyma [10], leading to systematic segmentation errors. Even readily available segmentations in widely used datasets, such as the developing Human Connectome Project (dHCP), contain inaccuracies in the segmentation of the CC, including ambiguous boundaries with adjacent structures (ventricles, ganglionic eminence, fornix) and discontinuities, largely due to directly adapting a neonatal segmentation approach [21] to the fetal context.

In postnatal and adult imaging, CC segmentation commonly exploits T1-weighted MRI or diffusion tensor imaging (DTI) [6]. DTI captures a defining property of the CC: its predominantly unidirectional fibers lead to high fractional anisotropy (FA) imaging values, while surrounding tissues display more isotropic values. In color FA (cFA) maps, these fibers are encoded in red under canonical orientation, forming a potential basis for unsupervised segmentation approaches. Although technically challenging, *in utero* DTI is feasible and increasingly explored for pathological assessment [4, 7, 19, 23], yet automated CC segmentation tools for fetal DTI remain scarce.

In this work, we propose a multi-modal segmentation strategy within an nnU-Net framework [12] with three objectives: (i) constructing a large-scale dataset using FA and cFA initialization to generate high-fidelity ground-truth annotations; (ii) introducing single- and dual-channel architectures integrating T2w and FA; and (iii) performing a comprehensive spatiotemporal characterization of CC development through morphometric (volume, length, regional thickness) and microstructural (FA profiles) features, enabling the definition of normative trajectories that can serve as a basis for detecting developmental deviations.

2 Materials and methods

2.1 Dataset and preprocessing

We included 192 fetal subjects from the dHCP cohort, ranging from 21 to 38 gestational weeks [11]. The dataset includes reconstructed T2w images ($0.5 \times 0.5 \times 0.5$ mm^3) [14] realigned to a common space and preprocessed dMRI data (15 $b = 0\,\mathrm{s/mm^2}$, 46 $b = 400\,\mathrm{s/mm^2}$, 80 $b = 1000\,\mathrm{s/mm^2}$) at 2 mm isotropic resolution [2]. From the overall cohort of 250 subjects, we selected those with both T2w and dMRI data ($n = 239$). The reconstructed T2w and dMRI volumes were then assessed for quality by an expert radiologist through visual inspection, resulting in the exclusion of subjects with insufficient image quality. The remaining dMRI volumes were coregistered to the T2w space using the rigid transformation matrices provided by dHCP [18]. Linear interpolation was applied, and all diffusion data were resampled onto the higher-resolution T2w grid. As an additional quality-control step, the accuracy of the T2w–dMRI alignment was visually assessed. Overall, these quality-control procedures resulted in the exclusion of 47 subjects. Subsequently, diffusion tensor-derived maps, including FA and cFA, were computed using the REKINDLE algorithm [26] implemented in ExploreDTI [17], which provides robust tensor fitting to reduce the influence of noise and outliers (Fig. 1).

2.2 Corpus callosum annotation protocol

Anatomically, the CC was defined as the coherently organized midline white matter tract connecting the two cerebral hemispheres, distinguished by consistent

[11] dHCP: Approved by the UK Health Research Authority (REC: 14/LO/1169), with written parental consent. https://biomedia.github.io/dHCP-release-notes/

fiber alignment along its rostro–caudal axis. As such, CC raw 3D segmentation masks were generated using FA and cFA maps following a four-step procedure. First, to reduce intensity peaks at the cortical plate, each FA map was eroded and then thresholded at the 85^{th} percentile. Second, from the cFA map we extracted candidate CC voxels by identifying regions where the red channel, corresponding to the left–right fiber orientation, is higher than green and blue components. Third, we selected the largest connected component to eliminate spurious clusters. Fourth, all automatically generated masks were manually refined with ITK–SNAP [29] by two trained raters under the supervision of an expert radiologist to correct common segmentation artifacts. Inter-rater reliability for CC volume and area measurements was evaluated on a subset of 17 subjects spanning the full gestational age (GA) range and resulted in excellent [13] agreement between raters (Intraclass Correlation Coefficients $> 95\%$; $p < 0.001$).

2.3 Automated corpus callosum segmentation

For the CC segmentation, we exploited the 3D nnU-Net framework [12], a self-configuring deep learning approach that automatically adapts network design, preprocessing and training strategies to input data. The dataset was split into training and testing ($77/23\%$) while preserving balance across GAs.

Three separate CC segmentation models were trained using different input configurations: T2w images only ($T2w$), FA images only (FA), and a dual-channel approach combining both T2w and FA ($T2w + FA$). Training strategy followed the standard nnU-Net pipeline with five-fold cross-validation on the training set. Final predictions were obtained using an ensemble of the five

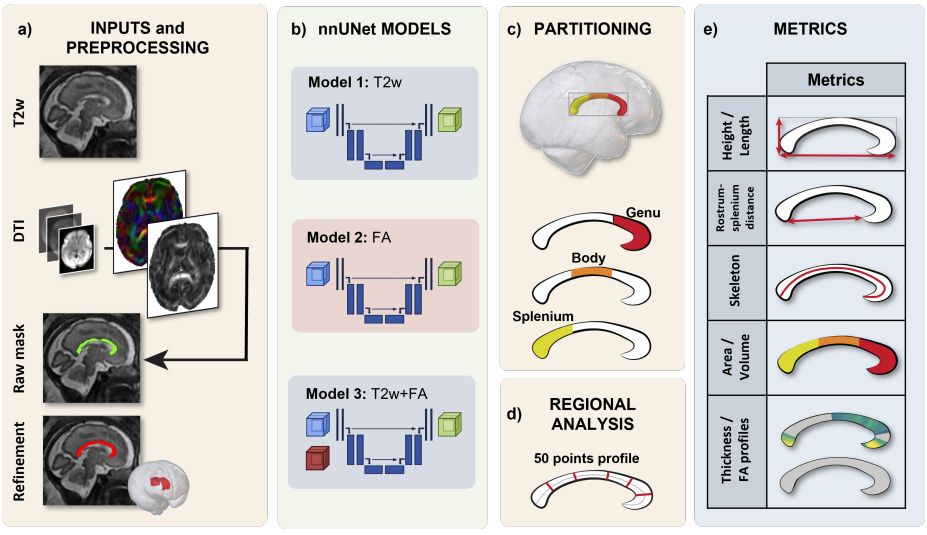

Fig. 1. Flow chart illustrating the processing pipeline for CC characterization.

cross-validation folds, by averaging the outputs of the corresponding trained models. The resulting model was then evaluated exclusively on an independent test set that was not used during training or validation. Model performance was assessed using the following segmentation metrics: 3D Dice similarity coefficient (DSC) and 2D Dice similarity coefficient on midsagittal slices (DSC2D) to measure overlap, volume similarity (VOLSIM), 95^{th} percentile Hausdorff Distance (HD95, mm) to quantify boundary discrepancies, Euler Difference (ED) to assess topological consistency, and Normalized Surface Distance (NSD) to evaluate surface agreement within a defined tolerance. Together, the 2D and 3D metrics provided a comprehensive evaluation of segmentation accuracy and structural fidelity.

To assess the robustness of the models against variations in diffusion MRI protocols, we conducted an ablation study by undersampling the original dMRI acquisition. Specifically, the test set was downsampled to simulated protocols with 1 $b = 0\,\mathrm{s/mm^2}$ and 15 $b = 1000\,\mathrm{s/mm^2}$ (hereafter denoted as FA_{15}). This reduced protocol more closely reflects dMRI acquisitions commonly used in fetal context [8].

All models were trained on NVIDIA L40S GPU (46GB VRAM).

2.4 Spatiotemporal characterization of corpus callosum development

To characterize CC development across GAs, we extracted a comprehensive set of morphometric and microstructural features from the masks predicted by the best performing model for both training and testing subjects. The midsagittal slice was automatically identified as the plane with the maximum CC cross-sectional area, selected within a ± 2 slice neighbourhood.

An oriented bounding box was estimated using minimum-area rectangle fitting on the 2D midsagittal mask. The longest axis of the bounding box defined the principal CC axis and was used to partition the mask into three equal anatomical regions corresponding to the genu, body, and splenium (Fig. 1c) [28].

The following morphometric (MP) and microstructural (MS) metrics were extracted from the CC masks (Fig. 1e):

- *Height and length (MP)*: derived from the shortest and longest dimensions of the oriented bounding box, respectively.
- *Rostrum–splenium distance (MP)*: the outer contour of the CC was extracted on the midsagittal slice using the marching squares algorithm for edge detection. The rostrum and splenium were identified as points of maximum curvature in the anterior and posterior contour regions, and the Euclidean distance between them was calculated [25].
- *Skeleton (MP)*: the CC contour was divided into superior and inferior boundaries using the rostrum and splenium landmarks as connecting points. Subsequently, for each point on the superior boundary, the nearest point on the inferior boundary was identified, and their midpoint was used to identify the CC midline. The midline was resampled into 50 equally spaced points along

arc length (Fig. 1d) and its length was defined as *skeleton*. Unlike *length*, this metric accounts for the intrinsic curvature of the CC, providing a more anatomical characterization of its morphology.
- *Area (MP)*: computed for the whole CC and for each anatomical subdivision by averaging measurements across the midsagittal slice and adjacent slices within a ± 2 slice range.
- *Volume (MP)*: calculated for the whole CC and separately for the genu, body, and splenium subdivisions.
- *Thickness (MP)*: the local thickness for each midline point was defined as the length of the shortest straight line that passes through the midline point and connects the superior and inferior CC boundaries.
- *FA profile (MS)*: FA values were sampled at each of the 50 midline points on the midsagittal plane using nearest-neighbor interpolation, yielding a spatially resolved 50-point FA profile along the CC.

2.5 Statistical analysis

Statistical comparisons between models were performed through paired t-tests ($p < 0.01$, Bonferroni-corrected). Furthermore, the best performing model was evaluated on the test set against two state-of-the-art methods trained on T2w images: (i) fss-CC [11], a CC-specific segmentation method using pathology-informed domain randomization approach; and (ii) multi-BOUNTI [27], a model based on a 3D Attention U-Net trained on dHCP subjects.

The extracted metrics were modeled as functions of GA using linear, quadratic, and exponential models, with the optimal model selected based on the Akaike Information Criterion (AIC). To characterize CC regional variations in thickness and FA trajectories along the 50-point profiles, a general linear model was applied at each point with GA as the independent variable. Statistical significance of GA-related effects was assessed to identify regions exhibiting significant developmental changes ($p < 0.05$, FDR-corrected). Rates of change were quantified using the slope coefficient (β).

To assess agreement between automatic and manual segmentations, Spearman rank correlation was used to compare morphometric and microstructural features derived from network predictions versus manual annotations (ρ, $p < 0.05$, FDR-corrected).

3 Results and Discussion

3.1 Network segmentation performance

Both single and dual-channel nnU-Net segmentation models achieve good performance (DSC > 0.70, DSC2D > 0.75) and optimal topological consistency (ED $= 0$). The dual-channel ($T2w + FA$) model achieved numerically the highest performance across all evaluation metrics, with statistically significant improvements over the T2w-only model on almost every metric (DSC2D, $p = 0.015$), and over the FA-only model specifically in DSC, DSC2D, and NSD; differences in

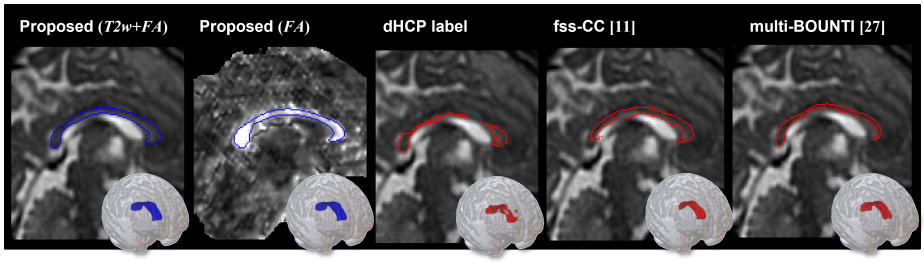

Fig. 2. Comparison of automated CC segmentation methods on a representative mid-sagittal T2w/FA slice of a fetus at 36 weeks.

volume similarity and HD95 relative to the FA-only model did not reach significance (Table 1). When benchmarked against state-of-the-art methods [11, 27], our dual-channel model exhibits improved accuracy (DSC and DSC2D > 0.7) and topological consistency (proposed, ED $= 0$ vs. fss-CC, ED $= 0.34$ and multi-BOUNTI, ED $= 0.5$), with higher improvement against dHCP label (ED $= 3.95$). Boundary agreement was also improved relative to both benchmarks: HD95 was lower (8.28 mm vs. 18.43 mm for fss-CC and 10.19 mm for multi-BOUNTI), while NSD exceeded 0.75, compared with 0.62 and 0.70, respectively. Qualitative assessment of CC segmentation by an expert radiologist reveals characteristic failure modes in the state-of-the-art methods: fss-CC [11] systematically failed to segment the splenium, likely a consequence of its agenesis-simulating augmentation strategy, while multi-BOUNTI [27] consistently under-segmented the genu, possibly due to the adaptation of tools originally developed in neonatal context [21]. In contrast, our model better captures the splenium and body, with minor over-segmentation in the genu (Fig. 2).

Results on the downsampled FA_{15} test set (Table 2) indicate that the $T2w+FA$ model maintains relatively robust performance despite the use of undersampled FA images, supporting its adoption in routine clinical acquisition protocols. In contrast, the FA model exhibits a marked performance drop, suggesting that

Table 1. Comparison of CC segmentation performances across models and statistical significance, i.e., p-values from t-tests.

Method	**DSC↑**	**VOLSIM↑**	**HD95↓**	**ED**	**DSC2D↑**	**NSD↑**
$T2w + FA$	$\mathbf{0.74 \pm 0.07}$	$\mathbf{0.92 \pm 0.06}$	$\mathbf{8.28 \pm 2.85}$	**0**	$\mathbf{0.79 \pm 0.08}$	$\mathbf{0.81 \pm 0.07}$
$T2w$	0.71 ± 0.07	0.90 ± 0.07	9.08 ± 3.13	**0**	0.78 ± 0.08	0.77 ± 0.08
FA	0.71 ± 0.08	0.92 ± 0.06	8.35 ± 2.65	**0**	0.76 ± 0.11	0.78 ± 0.07
fss-CC [11]	0.55 ± 0.08	0.81 ± 0.11	18.43 ± 17.89	0.34	0.70 ± 0.09	0.62 ± 0.08
multi-BOUNTI [27]	0.62 ± 0.08	0.83 ± 0.10	10.19 ± 3.09	0.50	0.72 ± 0.10	0.70 ± 0.08
p-values						
$T2w + FA$ vs $T2w$	< 0.001	< 0.005	< 0.001	1	0.015	< 0.001
$T2w + FA$ vs FA	< 0.001	0.409	0.730	1	0.001	0.001
$T2w + FA$ vs fss-CC	< 0.001	< 0.001	< 0.001	0.005	< 0.001	< 0.001
$T2w + FA$ vs multi-BOUNTI	< 0.001	< 0.001	< 0.001	0.007	< 0.001	< 0.001

segmentation based solely on FA images is highly sensitive to diffusion undersampling. Moreover, the combined $T2w + FA$ model continues to significantly outperform the $T2w$-only approach, in terms of DSC and NSD scores ($p < 0.001$).

3.2 Morphometric and microstructural features

Morphometric analysis revealed robust ($\rho > 0.5$) GA-dependent growth trajectories for all MP parameters except CC thickness with distinct regional dynamics. Global metrics such as CC length, body area, and body volume exhibited continuous, accelerated growth extending into late gestation. In contrast, height, genu area, and splenium volume demonstrated subtle plateaus after 32–34 weeks (Fig. 3a). These trajectories closely align with normative curves reported for length and area in the literature [15], and are consistent with sonographic observations of rapid linear growth during the second trimester followed by a relative deceleration in the third [22]. CC length demonstrated the lowest dispersion around the fitted developmental curve ($R^2 = 0.65$), indicating lower dispersion of individual measurements compared to area- ($R^2 = 0.37$) or volume- ($R^2 = 0.46$) based metrics. This reduced variance likely reflects the lower sensitivity of length measurements to slice positioning and the robustness of the oriented bounding-box approach for detecting the principal axis. Conventional prenatal charts primarily rely on CC length as the most accessible metric [20]; our findings suggest that CC length shows the most consistent association with GA among the evaluated metrics, supporting its use as a primary parameter for routine clinical growth monitoring.

Nevertheless, additional indices, such as thickness and skeleton, could provide superior quantitative phenotyping [9] and have shown strong alignment with ultrasound findings [15]. Regional analysis of thickness profiles revealed spatially heterogeneous growth rates along the CC, with the highest β slopes concentrated in the genu and splenium (Fig. 3b). In the presence of abnormally thin or thick phenotypes, integrating normative trajectories of global length with regional thickness, area, and volumetric measurements may substantially improve morphological characterization [24, 15, 11]. Beyond whole-structure analysis, we provided quantitative characterization of CC subcomponents, which could serve as sensitive biomarkers for subtle partial agenesis, moving clinical assessment beyond binary "present/absent" classifications [15, 11]. Contrary to other metrics, the FA profile remained remarkably stable (Fig. 3a). Although FA values showed greater absolute variability, no significant linear or non-linear temporal trends were detected. The analysis of profiles corroborated this, yielding no significant regions of age-related change. This pattern aligns with established neu-

Table 2. Results of $T2w + FA$ model on FA_{15} test set and statistical significance with respect to the original diffusion protocol (* $p < 0.01$).

Method	DSC↑	VOLSIM↑	HD95↓	ED	DSC2D↑	NSD↑
$T2w + FA$	0.73 ± 0.06*	0.90 ± 0.08	8.71 ± 2.94*	0	0.78 ± 0.08*	0.79 ± 0.07*
FA	0.64 ± 0.14*	0.87 ± 0.16*	12.05 ± 9.31*	0	0.68 ± 0.15*	0.69 ± 0.15*

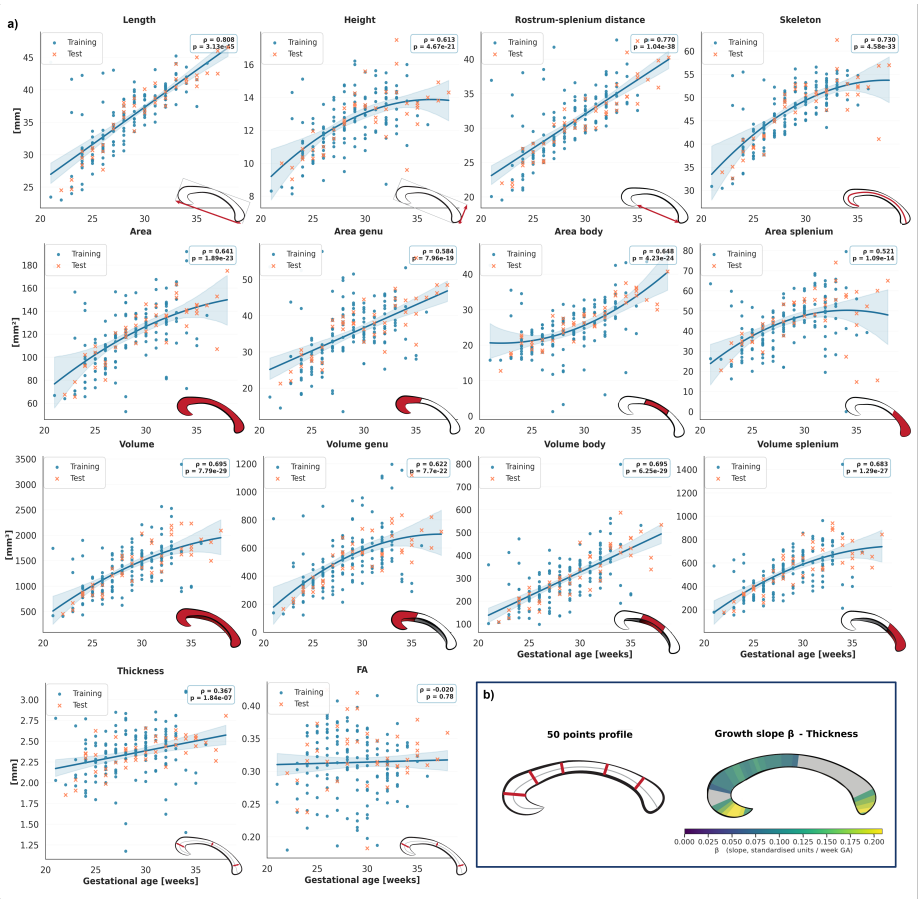

Fig. 3. a) Normal temporal evolution of CC metrics from 21 to 38 weeks. Metrics were derived from network predictions for 148 training subjects (dots) and 44 testing subjects (crosses), together with the best-fit model (solid line) and 95% bootstrap intervals (shaded region). b) Regional growth slope (β) mapped along the CC midline; warmer colors indicate faster thickness increase, while gray regions denote points where the slope coefficient (β) did not reach statistical significance ($p < 0.05$, FDR-corrected).

rodevelopmental models wherein early axonal organization, fasciculation, and pathway crossing precede true myelination [23, 4, 1]. Our findings confirm that pre-myelination axonal architecture dominates fetal white matter anisotropy, whereas myelination-driven FA acceleration remains a strictly postnatal event.

Finally, Spearman correlation between network predictions and manual annotations of morphometric and microstructural features confirmed strong concordance ($\rho > 0.77$, all $p < 0.001$), with the highest agreement for volumetric measures ($\rho > 0.93$) and FA profiles ($\rho > 0.87$).

3.3 Limitations and future directions

This study is not without limitations. First, FA upsampling from the native 2 mm to 0.5 mm resolution could introduce interpolation artifacts, particularly in thinner callosal regions; nonetheless, the combined $T2w + FA$ signal retained sufficient discriminative power for accurate boundary delineation and topological consistency (ED = 0). Furthermore, because ground-truth masks were initialized from FA and cFA maps prior to manual refinement, the FA-only and dual-channel models are not evaluated against a label set fully independent of their own input signal. This may partly explain their advantage over the T2w-only model, particularly along ambiguous boundaries where raters could have been implicitly guided by the FA-derived prior. The high inter-rater reliability (ICC $> 95\%$) and the requirement for expert-supervised correction of the automated initialization limit, but do not eliminate, this concern.

A second, related limitation is that our cohort was restricted to subjects with high-quality T2w and dMRI. Consistent with this, our undersampling experiment (FA_{15}) varied only the number of diffusion-weighted directions on an otherwise clean, high-resolution acquisition, and so does not capture the additional degradation introduced by thick slices or fetal motion artifacts typical of clinical diffusion protocols. At the same time, it is worth noting that under these high-quality conditions the FA-only model already matched the dual-channel model on volume similarity and boundary distance (1) while requiring no T2w input at all. This suggests that, once diffusion data are available, microstructural anisotropy on its own carries the information needed for accurate CC delineation.

The three models T2w-only, FA-only, and combined, can be seen as complementary options to be selected depending on which modalities are available and their relative quality, with cross-modal knowledge distillation as a natural next step for settings where dMRI cannot be acquired at all. Finally, all models were trained and evaluated on a single cohort (dHCP) restricted to typically developing fetuses. External validation across scanners, acquisition protocols, and pathological populations is needed; we view this as a promising direction for future work. In this view, the improved label quality could enable domain randomization strategies [11], and the resulting normative reference ranges could support quantitative assessment of neurodevelopmental deviations.

4 Conclusion

This work presents a multi-modal segmentation model that overcomes the limitations of T2w-only approaches by integrating microstructural anisotropy information. Beyond the dual-channel model ($T2w+FA$), we provide single-modality networks ($T2w$ and FA) trained on higher-quality ground-truth that can serve multiple purposes: reliable T2w-based segmentation when diffusion data are unavailable, and seeding for fetal tractography pipelines where automated CC delineation methods remain scarce.

Acknowledgments. This work was partially supported by grants from the Italian Ministry of Health to D. Peruzzo ("Ricerca corrente" funds). The research of

A. De Luca was supported by a Starting Grant from the European Research Council (agreement 101163214), a VIDI grant from the Dutch Scientific Council (NWO, Grant ID https://doi.org/10.61686/STMOQ50437), and by Stichting Hanarth Fonds. This work was also supported by the National Growth Fund AiNed Fellowship grant NGF.1607.22.006 to S.M.C. de Zwarte. The dHCP data were provided by the developing Human Connectome Project, KCL-Imperial-Oxford Consortium funded by the European Research Council under the European Union Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. [319456]. The financial support by the Austrian Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged. We are grateful to the families who generously supported this trial.

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

References

1. Calixto, C., et al.: Detailed anatomic segmentations of a fetal brain diffusion tensor imaging atlas between 23 and 30 weeks of gestation. Hum Brain Mapp. (2023)
2. Christiaens, D., et al.: Scattered slice shard reconstruction for motion correction in multi-shell diffusion mri. NeuroImage (2021)
3. Ciceri, T., et al.: Review on deep learning fetal brain segmentation from magnetic resonance images. Artif Intell Med. (2023)
4. Corroenne, R., et al.: Quantitative fetal mri with diffusion tensor imaging in cases with 'short' corpus callosum. Ultrasound Obstet Gynecol. (2024)
5. Corroenne, R., et al.: Prenatal evaluation, diagnosis and management of fetal corpus callosal abnormalities: international delphi consensus. Ultrasound Obstet Gynecol. (2025)
6. Cover, G.S., et al.: Computational methods for corpus callosum segmentation on mri: A systematic literature review. Comput Methods Programs Biomed. (2018)
7. Desai, S., et al.: Thick and short fetal corpus callosum on ultrasound: Added value of fetal magnetic resonance diffusion tensor imaging with tractography. Ped Neurol. (2025)
8. Di Stefano, M., et al.: Diffusion mri of the prenatal fetal brain: a methodological scoping review. NeuroImage (2025)
9. Garel, C., et al.: Biometry of the corpus callosum in children: Mr imaging reference data. AJNR Am J of Neurorad. (2011)
10. Gholipour, A., et al.: A normative spatiotemporal mri atlas of the fetal brain for automatic segmentation and analysis of early brain growth. Sci Rep. (2017)
11. Grifell I Plana, M., et al.: Enhancing corpus callosum segmentation in fetal mri via pathology-informed domain randomization. In: PIPPI. Springer (2026)
12. Isensee, F., et al.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods (2021)
13. Koo, T.C., et al.: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. (2016)
14. Kuklisova-Murgasova, M., et al.: Reconstruction of fetal brain mri with intensity matching and complete outlier removal. Medical Image Analysis (2012)

15. Lamon, S., et al.: Assessment of fetal corpus callosum biometry by 3d super-resolution reconstructed t2-weighted magnetic resonance imaging. Front in Neurol. (2024)
16. Lanzarone, V., et al.: Fetal corpus callosum anomalies: A review of underlying genetic disorders and prenatal testing options. J Ultrasound Med. (2025)
17. Leemans, A., et al.: Exploredti: a graphical toolbox for processing, analyzing, and visualizing diffusion mr data. Proc Intl Soc Magn Reson Med. (2009)
18. Li, B., et al.: Insights into temporal and spatial dynamics of short association fiber formation in the human fetal brain. bioRxiv (2025)
19. Li, Y., et al.: Prenatal ultrasound reassessment of the corpus callosum based on cortical connectivity information. Ultrasound Med Biol. (2024)
20. Luis, A., et al.: Towards automated fetal brain biometry reporting for 3-dimensional t2-weighted 0.55-3t magnetic resonance imaging at 20-40 weeks gestational age range. Pediatric Radiology (2026)
21. Makropoulos, A., et al.: Automatic whole brain mri segmentation of the developing neonatal brain. IEEE TMI (2014)
22. Malinger, G., et al.: The corpus callosum: normal fetal development as shown by transvaginal sonography. American Journal of Roentgenology (1993)
23. Millischer, A.E., et al.: Feasibility and added value of fetal dti tractography in the evaluation of an isolated short corpus callosum: Preliminary results. AJNR Am J of Neurorad. (2022)
24. Pashaj, S., et al.: Biometry of the fetal corpus callosum by three-dimensional ultrasound. Ultrasound Obstet Gynecol. (2013)
25. Peruzzo, D., et al.: A framework for the automatic detection and characterization of brain malformations: Validation on the corpus callosum. Med Image Anal. (2016)
26. Tax, C.M., et al.: Rekindle: robust extraction of kurtosis indices with linear estimation. Magn Reson Med. (2015)
27. Uus, A., et al.: Multi-bounti: Multi-lobe brain volumetry and segmentation for fetal and neonatal mri. medRxiv (2026)
28. Witelson, S.F.: Hand and sex differences in the isthmus and genu of the human corpus callosum. a postmortem morphological study. Brain (1989)
29. Yushkevich, P.A., et al.: User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage (2006)