scientific **data**

OPEN

DATA DESCRIPTOR

A benchmark for 2D foetal brain ultrasound analysis

Mariano Cabezas [1,4 ✉], Yago Diez[2,4], Clara Martinez-Diago[3,4] & Anna Maroto[3,4]

Brain development involves a sequence of structural changes from early stages of the embryo until several months after birth. Currently, ultrasound is the established technique for screening due to its ability to acquire dynamic images in real-time without radiation and to its cost-efficiency. However, identifying abnormalities remains challenging due to the difficulty in interpreting foetal brain images. In this work we present a set of 104 2D foetal brain ultrasound images acquired during the 20th week of gestation that have been co-registered to a common space from a rough skull segmentation. The images are provided both on the original space and template space centred on the ellipses of all the subjects. Furthermore, the images have been annotated to highlight landmark points from structures of interest to analyse brain development. Both the final atlas template with probabilistic maps and the original images can be used to develop new segmentation techniques, test registration approaches for foetal brain ultrasound, extend our work to longitudinal datasets and to detect anomalies in new images.

Background & Summary

Foetal brain development is a complex sequence of events ocurrying throughout gestation. From early stages of embryonic development, the brain undergoes structural changes until several months after birth[1,2]. Therefore, understanding normal brain development is essential to identify potential deviations which may lead to neurological disability. Specifically, various studies have described normal milestones within a specific chronology[3–5]. These changes can be observed by experts in prenatal diagnosis with skills in foetal neurology using ultrasound (US) and magnetic resonance imaging (MRI). However, identifying neurodevelopmental deviations is challenging due to the difficulty in image interpretation[5].

Currently, US is the established technique for screening due to its ability to acquire dynamic images in real-time without radiation and cost-efficiently. The international guidelines recommend the acquisition of a routine first trimester US within the 11th - 14th weeks and a mid-trimester scan within the 19th - 22nd weeks of gestation for anatomical evaluation of the foetus[6]. The standard foetal assessment includes the evaluation of planes acquired using 2D-US. Even though 3D-US is considered useful for prenatal diagnosis of some disorders (mainly involving the face, the skeleton, the cardiovascular system or the brain), the major obstacles for 3D-US implementation worldwide as the main routine acquisition type are related to foetal motion artefacts and acoustic shadowing during volume acquisition. Another widely used prenatal imaging technique is MRI. However, MRI is not used as a primary screening tool and is instead used as a complementary acquisition when foetal abnormalities are suspected (even in selected high-risk cases)[7,8]. Moreover, foetal MRI performed before the 18th-22nd weeks does not usually provide additional information to that obtained by US. Generally, MRI provides a detailed visualization of structures between the 26th and 32nd weeks, being superior to US and less susceptible to limitations from maternal body conditions and foetal presentation (bones do not produce occlusion artefacts in MRI)[8]. Within this context, a foetal brain US atlas based on 3D scans of healthy foetuses was recently published[9]. This atlas demonstrates the feasibility of assessing structural changes in the cortex and in the subcortical grey matter by US. Furthermore, this chronological US-based atlas of the foetal brain at each gestational age has the potential to become a useful tool to detect development abnormalities when combined with the standard planes of mid-trimester routine scans. However, considering that the primary diagnostic tool in pregnancy remains 2D-US, studies aimed at improving the identification of fetal structures should focus on 2D-US.

[1]Brain and Mind Centre, University of Sydney, Sydney, Australia. [2]Faculty Of Science, Yamagata University, Yamagata, Japan. [3]Hospital Universitari de Girona Doctor Josep Trueta, Girona, Spain. [4]These authors contributed equally: Mariano Cabezas, Yago Diez, Clara Martinez-Diago, Anna Maroto. ✉e-mail: mariano.cabezas@sydney.edu.au

(a) **Left**: Automatic Skull segmentation (Unet) example. **Right**: Ellipse fitted to the segmentation with major, minor axes, its centre and annotated landmarks highlighted.

(b) **Left**: image transformed after registration. **Right**: Reference (fixed) image centred.
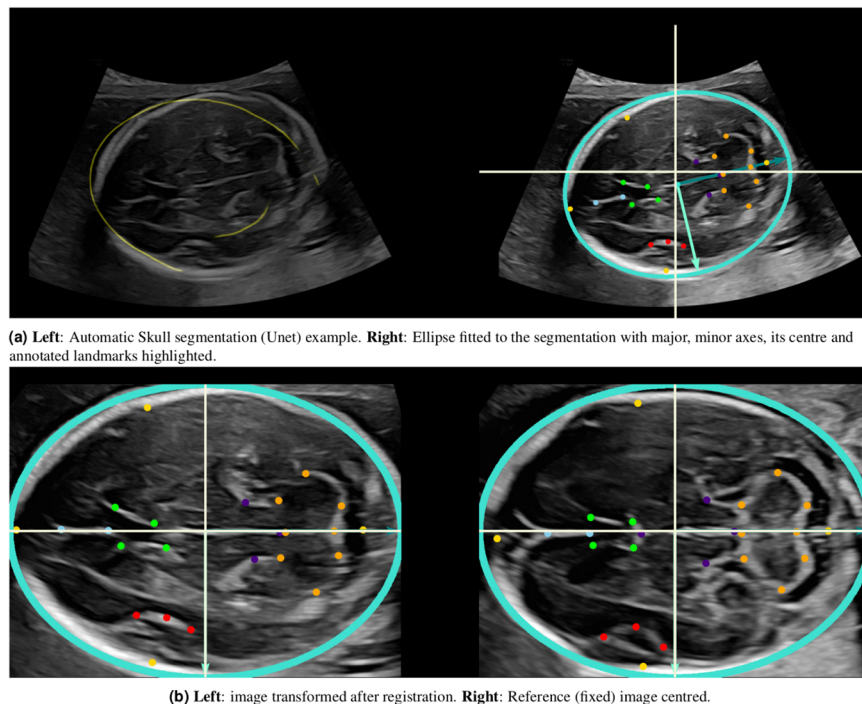
Fig. 1 Methodology: 1) The skull is automatically segmented using a UNet network (top-left). 2) An ellipse is fitted to the skull segmentation (top-right) 3) to estimate an affine transformation to a reference image (bottom-right). The axes of the fitted ellipse are warped to the image coordinate axes and are re-scaled to fit the ellipse in the reference image (bottom-left).

As there is a degree of variability in the shapes of anatomical regions between individuals, atlases are typically built by taking into account a number of images considered to fall within normal parameters. A crucial step in this process is finding corresponding regions between different images and warping them to a common space, a process known as image registration. In the case of US atlases, the lower signal-to-noise ratio (SNR), the differences in location for the foetuses due to the lack of a robust localization technique and the presence of image artefacts that may blur or shadow salient features and structure boundaries makes this process particularly challenging.

In this paper we present a set of foetal brain 2D-US images with manual landmark annotations of structures of interest and soft probabilistic maps for those structures based on these landmarks in a common space. The aim of this dataset is to provide a set of US images as a starting point to study registration and segmentation of foetal brain scans and provide tools for researchers focusing on foetal brain development. Furthermore, in combination with the recently published 4D atlas, this dataset can be also used to study 2D to 3D US registration and techniques to determine the "real age" of gestation and potential abnormalities when comparing to the atlas. To define a common space, a registration algorithm based on fitting an ellipse to the skull segmentation of a convolutional neural network (CNN) was used to align all the images to a common space. In order to avoid misalignments caused by image quality, a subject ellipse was used to estimate an affine transformation to a common (atlas) space. We also took advantage of the automatic differentiation capabilities of the pytorch package in three fundamentally different optimization settings (training a CNN for segmentation, fitting an ellipse to a segmentation boundary, and estimating an affine transformation).

Methods

In order to create the dataset presented, the data collected was processed as follows: First, the skull was automatically segmented using a UNet network, then an ellipse was automatically fitted to the shape of the skull. These ellipses where used to register each image to a reference image. Soft probabilistic maps where built using the set of registered images. The rest of this section includes details on each of these steps and a summary of other publicly available datasets. A visual summary can be found in Fig. 1.

Ultrasound data. A prospective cohort of low-risk pregnant women was recruited at routine mid-trimester foetal ultrasound scan. All participants initiated antenatal care before the 12th weeks of gestation, underwent the first trimester ultrasound scan between 12th and 14th weeks and had a low risk for aneuploidies in the first trimester combined screening. Written informed consents were obtained from participants. A private dataset of 70 pregnant women with a routine mid-trimester foetal ultrasound scan at [20 ~ 20.6] weeks without detected abnormalities was acquired, totalling 104 scans (8 women were scanned three times, 18 women were scanned twice and the remaining 44 women were only scanned once). The median of the maternal age was 31 (range 18-42). Images were acquired using high-frequency transabdominal probe (C2-9) of Voluson E10 ultrasound

system. For each subject, a transverse view of the foetal head demonstrating a standard normative transcerebellar scanning plane was manually annotated by 2 trained clinical experts (Fig. 1 top right). Both experts were part of the Prenatal Diagnosis Unit, one with over 10 years of experience and the other with 5 years of experience. The experts jointly participated in the annotations of each image to highlight common structures related to brain development as follows:

- **Skull**: 4 landmarks from the inner line of the skull were annotated: 2 at the level of the middle line and the other 2 in a perpendicular imagined line at the level of the posterior corners of the *cavum septi pellucidi* (CSP).
- **Cerebellar peduncles (Thalami)**: 1 landmark marking the edges of both thalami at the middle line and the outer edges of the concavity shape were annotated (3 total points).
- **Cerebellum**: 8 landmarks on the perimeter the cerebellum were annotated. Specifically, 2 points from the midline, 2 points from the cerebellum external edges and 4 points in the middle of each cerebellum hemisphere.
- **Cavum septi pellucidi (Cavum)**: 4 points, each marking one corner of its rectangular shape, were annotated.
- **Sylvian Fissure (Sylvius)**: 2 landmarks, one for each sylvian fissure edge, and 1 landmark in the inflection point of the fissure were annotated. For all the images, only the inferior fissure was visible.
- **Midline**: 1 landmark in the upper edge of the midline and 1 landmark in the upper edge of the CSP were annotated.

An important aspect of acquisition is that the sylvian fissure was always scanned on the lower part of the image, irrespective of the head orientation (left to right or right to left). This phenomena has important implications for registration. When aligning all images to a common 3D space (for example a 3D atlas template) in order to have all images facing the same direction, the transformation can be modeled with 180° rotations over the y axis. If the transformation is limited to the 2D space (image coordinates), a mirroring operation has to be applied to align all images. In the code we provide to process the images, the second option is used for simplicity. Furthermore, our common space is oriented from left to right (anterior to posterior).

Data acquisition was approved by the ethics committee "Comitè d'Ètica d'Investigació amb Medicaments CEIM GIRONA" with the code 2023.067. Furthermore, the subjects were informed and consented to the open publication of the data.

Skull segmentation database. The segmentation dataset used to train a skull segmentation CNN was downloaded from the HC18 grand challenge on "automated measurement of foetal head circumference using 2D ultrasound images"[10,11]. The challenge comprises a set of 800×540-pixel 2D US images with a pixel sizes $\in (0.052, 0.326)$ mm. The data set was split into 999 images for training and 335 for testing. For each image in the training set, an ellipse was manually fitted to the HC by a trained sonographer but precise segmentations of the true skull boundary were not provided. No pre-processing techniques were applied to the images apart from computing the z-score of the intensity values with respect to the mean and standard deviation of the whole image (non-0 intensities) before feeding them to the segmentation network. The network trained with all the training set images was then used to roughly segment the skull of our own 2D US images and provide a starting point to fit an ellipse.

Other public databases. To the best of our knowledge, only 3 other large datasets focusing on 2-D ultrasound have been made publicly available, including HC18. As mentioned on our brief description of HC18, the annotations are limited (only an ellipse roughly representing the skull is given), the subjects present a large distribution of gestational ages and no further information on the acquired plan e is provided. For studies focusing on the segmentation of structures of interest or registration, new annotations would be needed. Furthermore, the gestational age and acquisition plane for each subject can have a large impact on the appearance of the structures. Another public dataset presented by Burgos-Artizzu et al.[12] includes a large set of images (12400) ranging from the 18th to the 40th week of pregnancy. Images were acquired using six different machines and labelled at the image level. Similarly to HC18, the dataset presents a large range of gestational ages and limited annotations that could only be used to develop classification algorithms. Finally, Alzubaidi et al.[13] released a public dataset of 3832 high resolution images. In contrast with the other two datasets, no mention of gestational age is provided and once again rough annotations in the form of bounding boxes are provided. Moreover, the authors also highlight image resampling as an additional shortcoming of their dataset.

In our dataset, we focus on a specific acquisition plane and gestational age as defined by international guidelines and raw images are provided. Furthermore, landmarks for the most salient points of each structure of interest are provided with software tools to estimate finer-grained masks and bounding boxes around the landmarks. In that sense, our dataset provides a useful tool to address multiple image analysis problems including registration that could not be easily approached with other available datasets and no additional annotations.

Registration method. Medical image registration[14,15] is a necessary initial step for many medical image processing applications that rely on group-wise analysis. Typically, images are registered in pairs: one of them is defined as the "fixed" image (or reference) and the other as the "moving" image. The moving image is then warped using a transformation function to generate the final "moved" image. The transformation is commonly optimised using a predefined metric that computes image similarity between the reference and the moved image after transformation. Here, we present a coarse registration method to roughly align different foetal ultrasound images as described in the following sections and illustrated in Fig. 1. We chose one of the images as the common "reference" image (image 10) and registered the remaining images to it.

Automatic Skull segmentation using a Unet. The Unet architecture[16] is one of the most common CNN architectures for image segmentation. Due to its encoder-decoder structure and the use of skip connections, precise segmentations based on multi-scale features can be obtained for a variety of segmentation applications[17]. Consequently, we used a 2D Unet trained on the HC18 challenge dataset to provide a rough segmentation of the skull for all the images in our registration dataset (see Fig. 1 for an example). Specifically, the encoder and decoder blocks are comprised of 6 convolutional layers with a residual connection[18] of 32, 32, 128, 256, 256 and 1024 features each (inverse order for the decoder) and a bottleneck of 1024 features. To optimise the weights of the Unet, the Adam algorithm with default initial learning rate was used to minimise the binary cross-entropy loss.

Ellipse Registration. An ellipse is a planar curve representing the locus of the points with constant added distances to two "focal points", as expressed by the quadratic equation:

$$f^E(X, Y) = AX^2 + BXY + CY^2 + DX + EY + F = 0, (X, Y) \in \mathscr{R}^2 \tag{1}$$

With points that satisfy $f^E(X, Y) \neq 0$ being inside the ellipse perimeter ($f^E(X, Y) < 0$) or outside of it ($f^E(X, Y) > 0$). The general equation's coefficients can be obtained from known semi-major axis a (represented by the magnitude of the turquoise vector in Fig. 1), semi-minor axis b (represented by the magnitude of the cyan vector in Fig. 1), centre coordinates (x_0, y_0) (represented by the point where the two vectors meet in Fig. 1) and rotation angle θ (the angle from the positive horizontal axis to the ellipse's major axis as observed in Fig. 1) using the formulae:

$$A = a^2\sin^2\theta + b^2\cos^2\theta \tag{2}$$

$$B = 2(b^2 - a^2)\sin\theta\cos\theta \tag{3}$$

$$C = a^2\cos^2\theta + b\sin^2\theta \tag{4}$$

$$D = -2Ax_0 - By_0 \tag{5}$$

$$E = -Bx_0 - 2Cy_0 \tag{6}$$

$$F = Ax_0^2 + Bx_0y_0 + Cy^2 - a^2b^2 \tag{7}$$

These expressions can be derived from the canonical equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ by an affine transformation of the coordinates (x, y) (with a translation $(-x_0, -y_0)$ and an angle θ).

$$x = (X - x_0)\cos\theta + (Y - y_0)\sin\theta \tag{8}$$

$$y = -(X - x_0)\sin\theta + (Y - y_0)\cos\theta \tag{9}$$

Given the set of pixel coordinates of the skull segmentation $((X_{skull}, Y_{skull}) = [(x_1, y_1), \ldots, (x_N, y_N)])$ we can fit an ellipse using its parameters (a, b, x_0, y_0 and *theta*) by minimising the following objective function:

$$\hat{a}, \hat{b}, \hat{x}_0, \hat{y}_0, \hat{\theta} = \arg\min_{a,b,x_0,y_0,\theta} \sum_{i=1}^{N} f^E(x_i, y_i)^2, \tag{10}$$

where the coefficients A to F are substituted in f^E by their definitions in equations (2) to (7) and the estimated ellipse parameters.

This process is repeated 5 times, removing erroneous points of the skull segmentation mask that are one standard deviation away from the mean ellipse error according to Eq. (10). With the parameters of the skull ellipse estimated through optimisation, we can now define an affine transformation (referred to as **Ellipse** from now on) to move the brain to the centre of the image as:

$$\alpha_e = \begin{bmatrix} \frac{w}{2a}\cos\theta & \frac{w}{2a}\sin\theta & -w/2 \\ -\frac{h}{2b}\sin\theta & \frac{h}{2b}\cos\theta & -h/2 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{11}$$

Affine image registration. For comparison, a regular rigid registration of 6 unrestricted parameters was performed with different initialisations. From the most basic identity initialisation (referred to as **Affine**), to an initialisation using the ellipse parameters of the reference image (referred to as **Affine (Reference ellipse)**) and a refinement of the ellipse-based affine transformation from Eq. (11) (referred to as **Ellipse + Affine**).
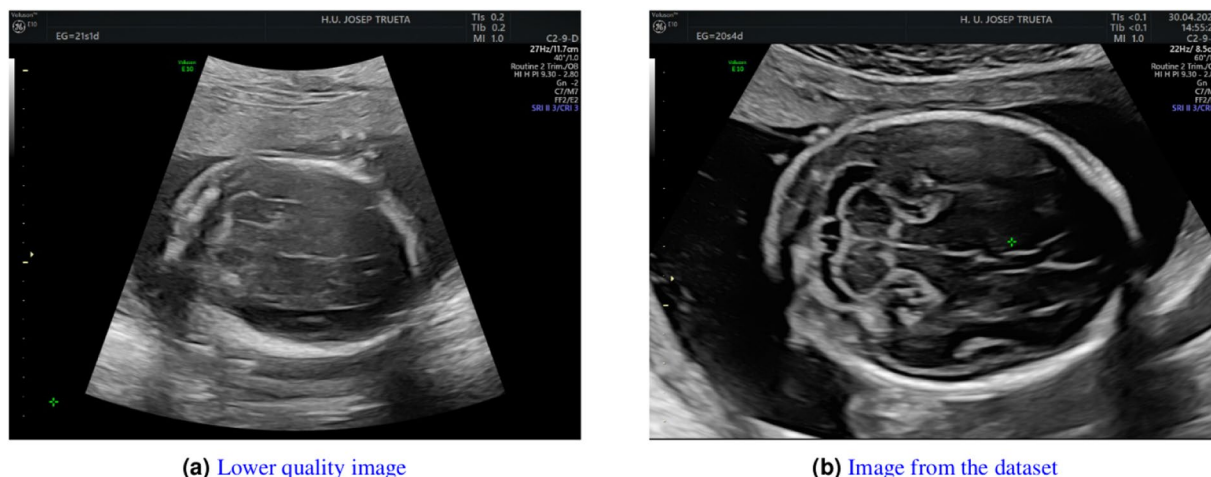
(a) Lower quality image   **(b)** Image from the dataset

Fig. 2 Qualitative comparison between a low quality image where structures are not clearly visible and and image from the dataset.

Probabilistic maps. Once the images are co-registered to a common space based on their ellipse, a two-step process is performed to generate probabilistic maps for all the structures with more than 2 landmarks.

First, the concave hull of each structure is computed using the alphashapes package. This concave hull represents a rough segmentation of the location of the structure. Second, the segmentations for all the subjects are averaged per structure to determine the probability of each pixel to belong to that structure.

This approach has some limitations. Namely, some structures have a polygonal shape, even though they are actually curves (e.g. sylvius) and the final masks are only a rough representation of the real boundaries (e.g. cerebellum). However, these polygonal maps can be still used to determine growth milestones and to provide a rough location of the structure of interest.

Data Records

The original images with manual landmark annotations (Gimp image editing tool format) and the co-registered images and probabilistic maps used within this paper can be found on figshare[19], and are organised with subject id number (1-52) and an additional number for multiple scans (e.g. 36 and 36.1). Co-registered images are saved in JPEG format with the "_registered" suffix. The final probability maps estimated using a combination of co-registered landmarks and the alphashape package are provided as JPEG images with the name of the structure (e.g. sylvius.jpeg), while comma separated value (CSV) files with the point landmarks of all registered subjects are compressed into a single zip file.

In total, the released dataset consists of 104 annotated 2D US images of foetal brains on the 20th week of pregnancy. The manual annotations are described in **Ultrasound dataset** section.

Technical Validation

To validate the techniques used for co-registration to a common space, we focused on common medical imaging metrics for registration using landmarks. In this section we describe these metrics and provide some qualitative and quantitative results of the alignment (including a visualisation of the probabilistic maps of one of the structures of interest).

Regarding the quality of the US images, all images were acquired with a high frequency ultrasound probe (2-9 MHz) that provides high resolution images following the ISUOG recommendations. The most constraining factor for the quality of the ultrasound images is maternal obesity. However, for this study we excluded women with maternal morbid obesity (body mass index > 40) as it is one of the factors for high-risk pregnancy. To further illustrate that point, we provide a comparative example between an image from the dataset and a lower quality one in Fig. 2.

Evaluation metrics. For this study we chose to use the anatomical knowledge defined by expert annotations as the main directive to evaluate the quality of the registration and avoid focusing exclusively on common pixel-metrics that might be unreliable and disconnected from physical properties[20]. Specifically, we used the point annotations described in the **Ultrasound Data** section with two point-based metrics and one area-based metric for every anatomical structure with more than 2 points (all the structures except the midline). For completeness sake, we also included the structural similarity index metric (SSIM) pixel-based metric.

Point-based metrics. In order to penalise partial matches between anatomical structures, we considered the *Hausdorff distance* (d_H), that computes the worst possible Euclidean distance ($d(p_i, p_j)$) between two sets of points P_f ($|P_f| = N$) and P_m ($|P_m| = M$). We also considered the *average of the minimum Euclidean distances* (d_E) to express global similarity between structures defined by landmarks.
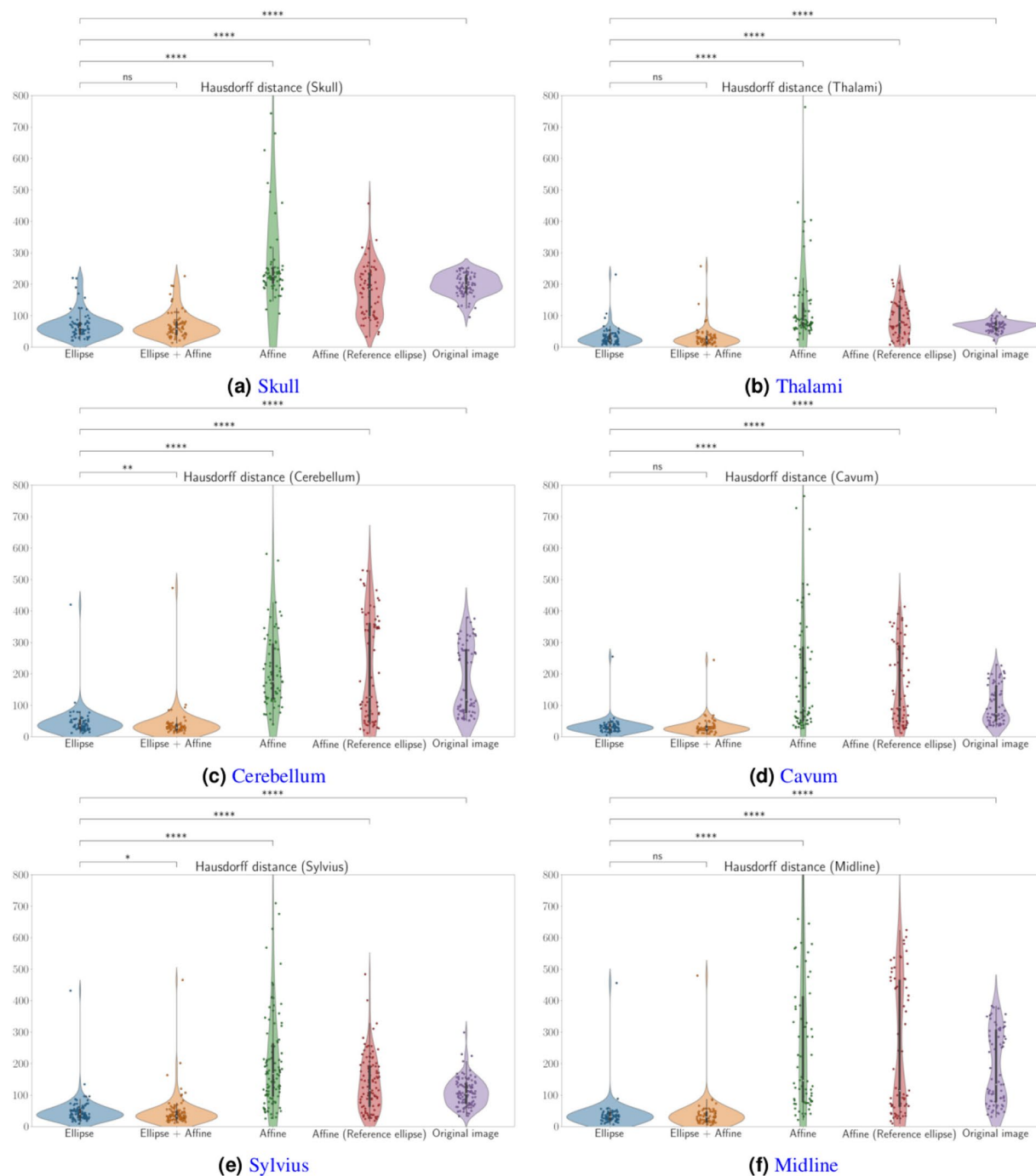
Fig. 3 Quantitative results for all the methods on each structure (Hausdorff distance, d_H, lower values indicate better registration). The upper part of each boxplot figure indicates the results of pairwise statistical Wilcoxon tests: (ns: 5.00e-02 $<$ p $\leq$ 1.00e+00, * 1.00e-02 $<$ p $\leq$ 5.00e-02, ** 1.00e-03 $<$ p $\leq$ 1.00e-02, *** 1.00e-04 $<$ p $\leq$ 1.00e-03, **** p $\leq$ 1.00e-04).

Area-based metric. To evaluate the superposition between two anatomical structures defined as 2D landmarks (points), we first computed their concave hulls and then considered their Dice similarity coefficient (polygon DSC).

Image similarity metric. For completeness, we also considered the structural similarity index measure (SSIM) as a pixel-intensity-based metric.

Comparison of coarse registration approaches. Figures 3, 4 and 5 summarise the results with boxplots and Wilcoxon signed-rank tests for all the registration methods and metrics considered. Wherever possible, the results for different anatomical areas are presented separately. For the Euclidean and Hausdorff metrics, lower
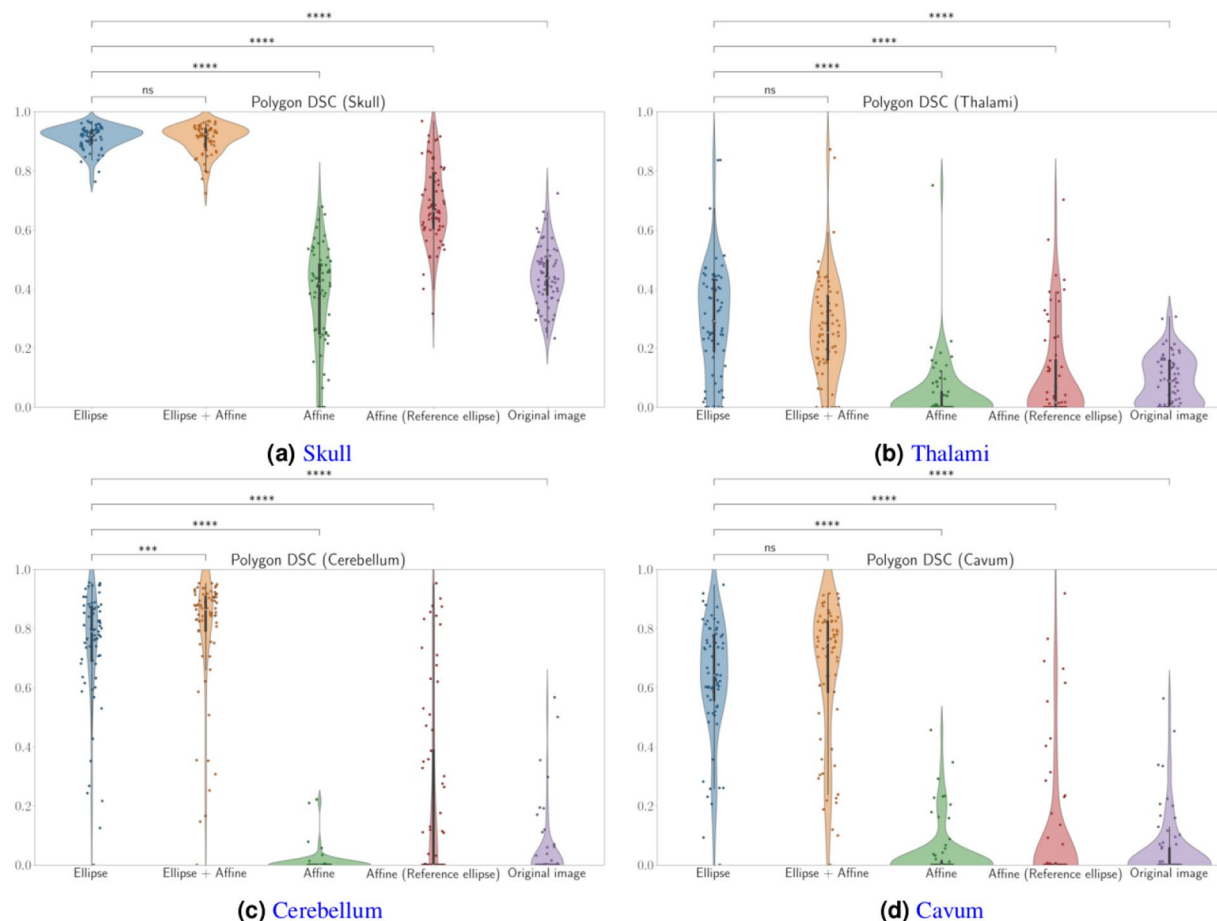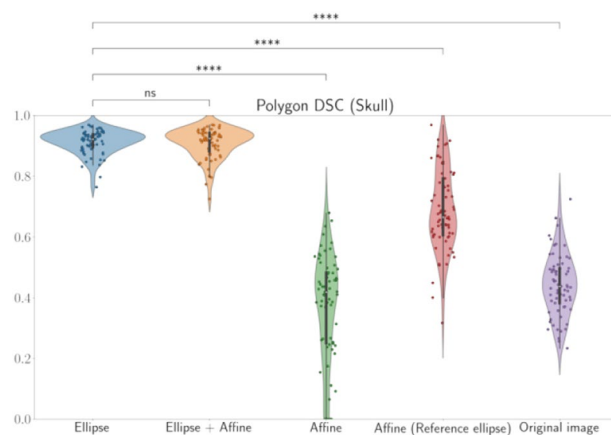
(a) Skull

(b) Thalami

(c) Cerebellum

(d) Cavum

Fig. 4 Quantitative results for all the methods on each structure with multiple points (polygon DSC, higher values indicate better registration). The upper part of each boxplot figure indicates the results of pairwise statistical Wilcoxon tests: (ns: 5.00e-02 $< p \leq$ 1.00e+00 * 1.00e-02 $< p \leq$ 5.00e-02, ** 1.00e-03 $< p \leq$ 1.00e-02, *** 1.00e-04 $< p \leq$ 1.00e-03, **** $p \leq$ 1.00e-04).

Fig. 5 Quantitative results for all the methods according to the SSIM metric (higher values indicate better registration). The upper part of each boxplot figure indicates the results of pairwise statistical Wilcoxon tests: (ns: 5.00e-02 $< p \leq$ 1.00e+00, * 1.00e-02 $< p \leq$ 5.00e-02, ** 1.00e-03 $< p \leq$ 1.00e-02, *** 1.00e-04 $< p \leq$ 1.00e-03, **** $p \leq$ 1.00e-04).

values indicate better registration, while for the SSIM and Dice metric higher results indicate better registration. Moreover, a qualitative example to illustrate misalignments between the reference points and the registered landmarks is provided in Fig. 6.
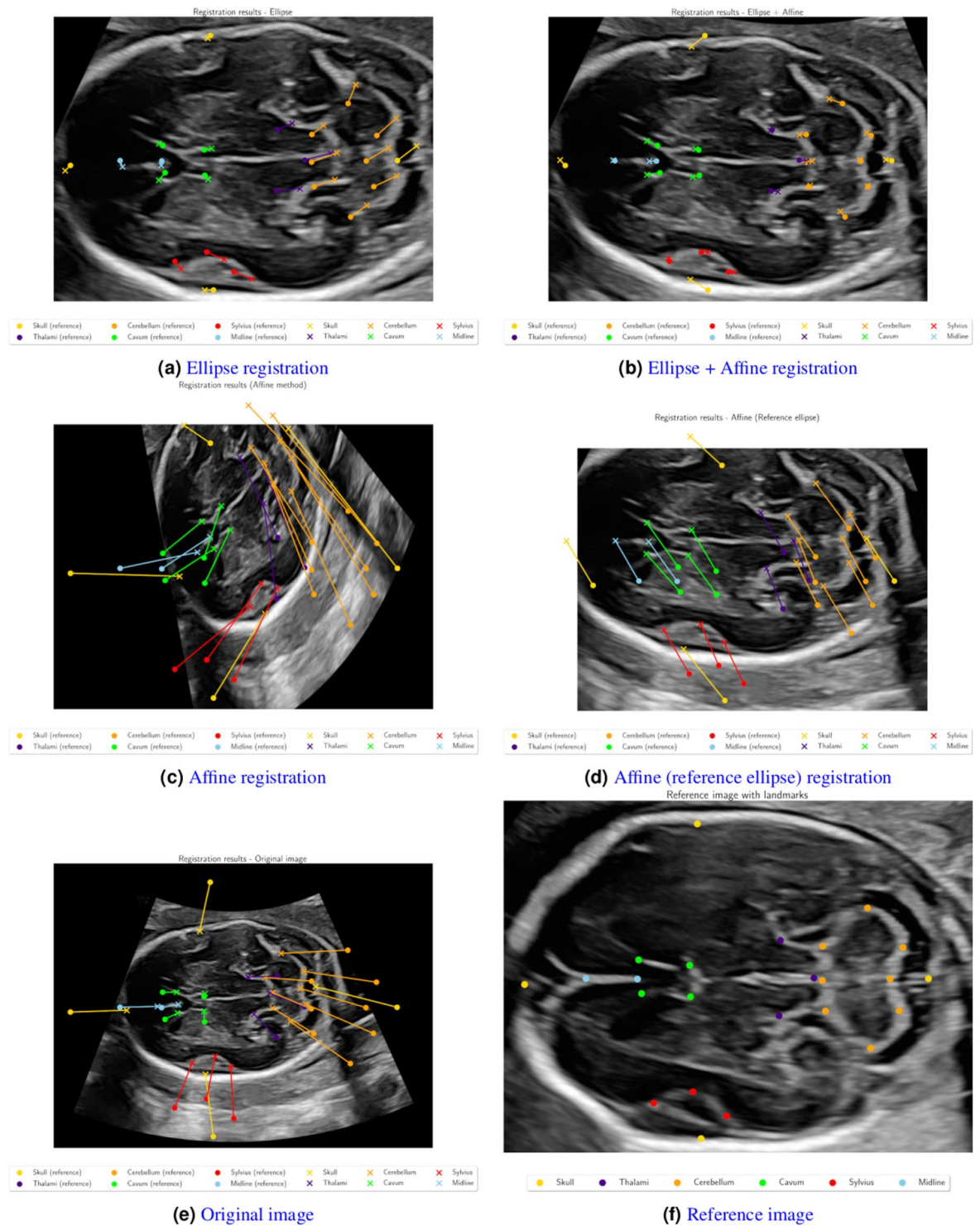
Fig. 6 Qualitative example of the alignment between the reference image and a randomly selected subject (11). The background image corresponds to the warped image after registration, circles denote the reference points for each structure, crosses represent the registered points and lines are used to illustrate misalignment between the reference and the registration.

Regarding point metrics, the ellipse (**E**) and ellipse with affine methods (**E+A**) obtain overall better results than the other methods. In general, the differences observed were found to be statistically significant for both point metrics and most anatomical structures. Exceptions to this are the thalami and the cerebellum where the pixel based affine registration method initialised using the reference ellipse (**AFF+I**), obtains results that appear worse but are not significantly different. Comparing the *E* and E+A methods, small (and not statistically significant) differences can be observed. Using a refinement registration after ellipse-based method slightly worsens
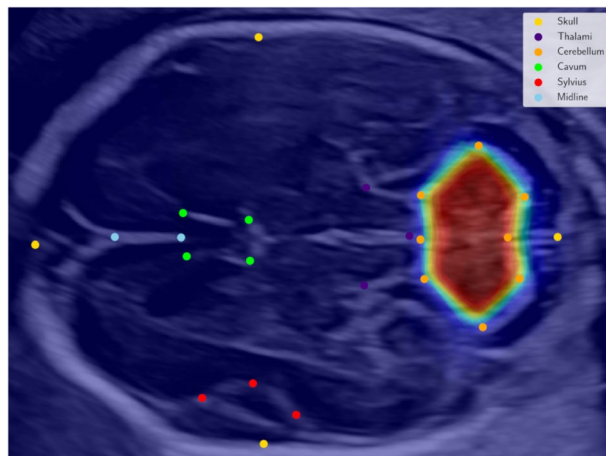
Fig. 7 Probabilistic map for the cerebellum structure based on the averaged concave hull of the landmarks for each image. Landmarks for the reference image (background) are also provided.

the metrics when applied to the skull but improves them in all other anatomical structures. This is an expected result as the main focus of the first method is to co-register the skulls (ellipses). The uninitialised affine transformation (**AFF**) obtains significantly worse results than these two methods and has a higher variance of values as illustrated by its wide boxplots. Both methods using exclusively pixel-wise affine registration (AFF, AFF+I) achieve results that are worse than the metrics of the original moving image in some cases. This illustrates the disconnect between pixel-based metrics guiding these methods and point-based metrics targeting distances between real anatomical structures (especially for noisy sequences).

Regarding the differences between the Euclidean and Hausdorff metrics (especially the median values shown by the central line in boxplots), slightly higher Hausdorff values indicate that some point pairs are further than the average euclidean distance mean value. This is especially noticeable on the skull for E and A+E and the thalami for *AFF*.

The other two metrics show the same general tendencies, even though they focus on different aspects of the registration. The low SSIM values observed for all methods (even though E and A+E outperform all methods) illustrats how challenging this registration scenario is. The variations between individuals and acquisitions and low SNR make the intensity values particularly unreliable. On the other hand, the high anatomical DSC results observed for the ellipse-based methods (includding AFF+I) validate our geometric approach that relies on a rough initial skull estimate to determine the general shape and orientation of the brain. Similarly high results obtained for the cerebellum indicate that a correct skull placement is a crucial important step towards the registration of all brain structures.

Finally, Fig. 7 shows the final heatmap generated from the registered concave hulls (E+A) of the cerebellum for all output images. Even though faint outlines of some incorrect registration results outside of the cerebellum region (delimited by orange points) can be observed, the higher probability regions in the heatmap clearly correspond to the cerebellum region of the reference image delimited by the manually annotated orange landmarks.

Usage Notes

The code provided to analyse the images and perform a rough initial alignment to a common space has been developed using python. A set of Jupyter notebooks detailing the use of each step is also provided in the repository with visualisation examples of each step of the processing pipeline. Furthermore, we provide the trained weights for the skull segmentation network as part of the data records (file *unet.pt*[19]).

Regarding the use of python packages, the code heavily relies on pytorch (version 1.12.0 with CUDA 10.2) to do the heavy lifting. However, numpy (version 1.21.6), scipy (version 1.7.3), scikit-learn (version 1.0.2), scikit-image (version 0.18.1), pandas (version 1.3.4), persim (version 0.3.1) and alphashape (version 1.3.1) are also used for some of the processing steps or to analyse different metrics related to the registration. Furthermore, opencv-python (version 4.5.2.52) and gimpformats (version 2021.1.4) where used to open the images with python. In particular, manual landmark annotation where done using the Gimp software and were stored (together with the ultrasound image) using the xcf format. Finally, matplotlib (version 3.7.1), seaborn (version 0.11.0) and statannot (version 0.2.3) were used for results visualisation inside the Jupyter notebooks.

Code availability

All the code used in the study to generate the final atlas and to co-register the images is publicly available at https://github.com/marianocabezas/fetal-brain.

References

1. Monteagudo, A. & Timor-Tritsch, I. E. Normal sonographic development of the central nervous system from the second trimester onwards using 2d, 3d and transvaginal sonography. *Prenatal Diagnosis* **29**, 326–339 (2009).
2. Namburete, A. I. *et al.* Learning-based prediction of gestational age from ultrasound images of the fetal brain. *Medical Image Analysis* **21**, 72–86 (2015).
3. Cohen-Sacher, B., Lerman-Sagie, T., Lev, D. & Malinger, G. Sonographic developmental milestones of the fetal cerebral cortex: a longitudinal study. *Ultrasound in Obstetrics & Gynecology* **27**, 494–502 (2006).
4. Kline-Fath, B. M. Ultrasound and mr imaging of the normal fetal brain. *Neuroimaging Clinics of North America* **29**, 339–356 (2019).
5. Guimaraes, C. V. & Dahmoush, H. M. Fetal brain anatomy. *Neuroimaging Clinics of North America* **32**, 663–681 (2022).
6. Salomon, L. J. *et al.* Isuog practice guidelines (updated): performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound in Obstetrics & Gynecology* **59**, 840–856 (2022).
7. Griffiths, P. D. *et al.* Mri in the diagnosis of fetal developmental brain abnormalities: the meridian diagnostic accuracy study. *Health Technology Assessment (Winchester, England)* **23**, 1 (2019).
8. Prayer, D. *et al.* ISUOG practice guidelines (updated): performance of fetal magnetic resonance imaging. *Ultrasound in Obstetrics & Gynecology* **61**, 278–287 (2023).
9. Namburete, A. I. *et al.* Normative spatiotemporal fetal brain maturation with satisfactory development at 2 years. Nature 1–9 (2023).
10. van den Heuvel, T. L. A., de Bruijn, D., de Korte, C. L. & van Ginneken, B. Automated measurement of fetal head circumference using 2d ultrasound images. *PloS one* **13**, e0200412 (2018).
11. van den Heuvel, T. L. A., de Bruijn, D., de Korte, C. L. & van Ginneken, B. Automated measurement of fetal head circumference using 2d ultrasound images [data set]. *Zenodo*, https://doi.org/10.5281/zenodo.1322001 (2024).
12. Burgos-Artizzu, X. P. *et al.* Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Scientific Reports* **10**, 10200 (2020).
13. Alzubaidi, M. *et al.* Large-scale annotation dataset for fetal head biometry in ultrasound images. *Data in Brief* **51**, 109708 (2023).
14. Fu, Y. *et al.* Deep learning in medical image registration: a review. *Physics in Medicine & Biology* **65**, 20TR01 (2020).
15. Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M. & Gopinath, K. Brain functional localization: A survey of image registration techniques. *IEEE Transactions on Medical Imaging* **26**, 427–451, https://doi.org/10.1109/TMI.2007.892508 (2007).
16. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 234–241 (2015).
17. Isensee, F., Jaeger, P. F., Kohl, S. A. A. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature: Methods (2021).
18. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *IEEE Conf. Comput. Vision Pattern Recog.*, 770–778 (2016).
19. Cabezas, M., Diez, Y., Martinez-Diago, C. & Maroto, A. A benchmark for 2d foetal brain ultrasound analysis [data set]. *Figshare*, https://doi.org/10.6084/m9.figshare.c.6984822.v1 (2024).
20. Rohlfing, T. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE Transactions on Medical Imaging* **31**, 153–163 (2012).

Author contributions

All authors conceived and designed the experiments, C.M. and A.M. collected the images and performed the landmark annotations, M.C. and Y.D. conducted the experiments and analysed the results. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.