Spatio-Temporal Atlas of Normal Fetal Craniofacial Feature Development and CNN-Based Ocular Biometry for Motion-Corrected Fetal MRI

Alena Uus[1(✉)], Jacqueline Matthew[1], Irina Grigorescu[1], Samuel Jupp[1], Lucilio Cordero Grande[2,3], Anthony Price[2], Emer Hughes[2], Prachi Patkee[2], Vanessa Kyriakopoulou[2], Robert Wright[1], Thomas Roberts[1], Jana Hutter[1], Maximilian Pietsch[2], Joseph V. Hajnal[1,2], A. David Edwards[2], Mary Ann Rutherford[2], and Maria Deprez[1]

[1] Biomedical Engineering Department, School of Imaging Sciences and Biomedical Engineering, King's College London, St. Thomas' Hospital, London, UK
alena.uus@kcl.ac.uk
[2] Centre for the Developing Brain, School Biomedical Engineering and Imaging Sciences, King's College London, St. Thomas' Hospital, London, UK
[3] Biomedical Image Technologies, ETSI Telecomunicacion, Universidad Politécnica de Madrid and CIBER-BBN, Madrid, Spain

Abstract. Motion-corrected fetal magnetic resonance imaging (MRI) is widely employed in large-scale fetal brain studies. However, the current processing pipelines and spatio-temporal atlases tend to omit craniofacial structures, which are known to be linked to genetic syndromes. In this work, we present the first spatio-temporal atlas of the fetal head that includes craniofacial features and covers 21 to 36 weeks gestational age range. Additionally, we propose a fully automated pipeline for fetal ocular biometry based on a 3D convolutional neural network (CNN). The extracted biometric indices are used for the growth trajectory analysis of changes in ocular metrics for 253 normal fetal subjects from the developing human connectome project (dHCP).

Keywords: Motion-corrected fetal MRI · Craniofacial features · Ocular measurements · Spatio-temporal atlas · Automated biometry

1 Introduction

Arguably, an MRI scan of the fetal brain is not complete without a structural and dysmorphological assessment of the fetal craniofacial structures due to the intricate link between brain anomalies and genetic syndromes that affect the facial features [2,19]. More than 250 syndromes are associated with changes in craniofacial growth and development and can therefore result in overt anomalies or subtle changes in anatomical appearance and yet prenatal detection remains

C. H. Sudre et al. (Eds.): UNSURE 2021/PIPPI 2021, LNCS 12959, pp. 168–178, 2021.
https://doi.org/10.1007/978-3-030-87735-4_16

low [7,15]. Ultrasound is the primary imaging modality for fetal assessment and has well recognised limitations. High risk fetal cases are increasingly referred for MRI examinations for further characterisation and to assess for the presence of additional anomalies.

Fetal motion during MRI acquisition leads to loss of structural continuity between 2D slices and corruption of 3D information. The image degradation precludes the reliable use of MRI to assess craniofacial structures. During the past decade this has been successfully addressed by slice-to-volume registration (SVR) reconstruction tools [6,11] that can produce motion-corrected high-resolution 3D fetal brain MRI images. SVR tools also have the potential to increase the clinical reliability of extended craniofacial biometry and objective assessments of the curved structures of the fetal head and face e.g. orbits, oral hard and soft palate, and cranial shape. Formalisation of the normal trajectory of rapid development of craniofacial structures occurring during gestation that can be observed in MRI is essential for definition of the control reference. However, the existing spatio-temporal fetal atlases include only the brain region [8].

Automated segmentation and volumetry methods lower the impact of inter- and intra-observer variability and provide the means for processing large-scale studies [14]. Recently, several reported works employed semi-automated segmentation for analysis of fetal craniofacial features in MRI and ultrasound [1,12,13,21] and, more recently, an automated method was proposed for 2D slice-wise segmentation and biometry of orbits in low resolution stacks [3]. Incorporation of novel convolutional neural network (CNN) pipelines for motion-corrected fetal MRI segmentation [9,10,16] has a potential to make the application of automated biometry and volumetry of craniofacial structures feasible for large datasets and motion-corrected MRI. Quality control for automated data analysis methods also remains one of the current challenges in terms of practical application.

In this work, we propose to generate a first spatio-temporal atlas of the fetal head that includes craniofacial features. This extends the already existing brain-only fetal MRI templates for a wider application of analysis of normal craniofacial feature development. In addition, we implemented an automated pipeline for 3D CNN-based ocular biometry for motion-corrected fetal MRI with outlier detection. The biometry outputs were then used for the analysis of ocular growth trajectories for 253 normal fetal subjects with acceptable biometry results.

2 Methods

2.1 Cohort, Datasets and Preprocessing

The data used in this study included T2w MRI datasets of 291 fetuses without reported anomalies from 20 to 38 weeks gestational age (GA) acquired at St.Thomas' hospital, London as a part of the dHCP project[1] (dHCP, REC: 14/Lo/1169). The acquisition was performed on a Philips Achieva 3T system

[1] dHCP project: http://www.developingconnectome.org/project.

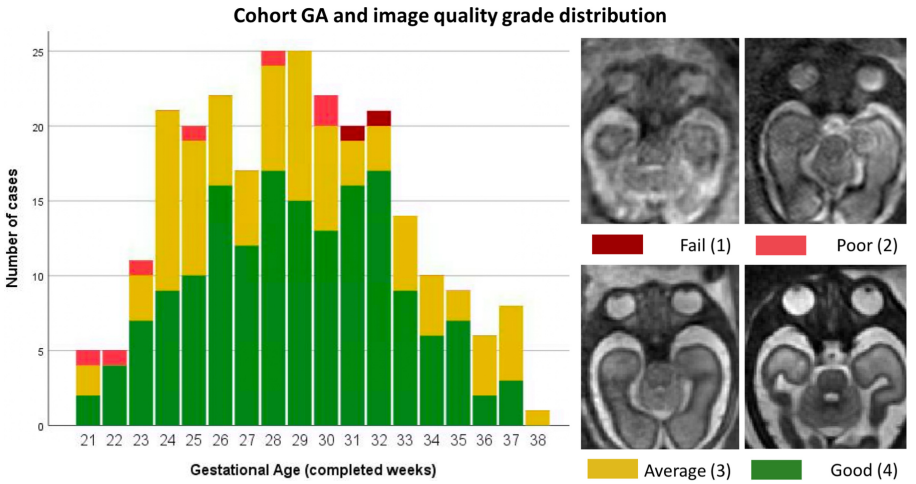

Fig. 1. Distribution of the gestational age and image quality in the investigated fetal MRI cohort. The image quality scores are: fail, poor, average and good.

with a 32-channel cardiac coil using single shot turbo spin echo (ssTSE) sequence with TE = 250 ms, TR = 2265 ms, acquisition resolution = $1.1 \times 1.1 \times 2.2$ mm (–1.1 mm gap) [18]. The datasets were reconstructed using a fully automated SVR pipeline [5] to $0.5 \times 0.5 \times 0.5$ mm resolution for the fetal head region of interest (ROI). This was followed by reorientation to the standard planes using a dedicated transformer CNN [22].

The quality of the 3D reconstructed images in terms of definition of the anatomy features, noise and contrast was assessed by an experienced researcher with the grades: good (4), average (3), poor (2) and failed (1). All available datasets were included in the biometry study irrespective of the reconstruction image quality for the purpose of testing of the proposed automated detection of outliers approach. The histograms of the cases GA and quality scores is given in Fig. 1 with the majority of scans within ≥ 3 quality window. For generation of the atlas we used only a subset of cases from 21 to 36 GA weeks.

2.2 Spatio-Temporal Atlas

For atlas generation, we selected 190 datasets with the best image quality and optimal coverage of the fetal head. The 4D spatio-temporal atlas of the fetal head was constructed using the MIRTK[2] atlas generation pipeline [20] at 16 discrete timepoints in 21 to 36 weeks GA range. We used local normalised cross-correlation similarity metric with 5 voxel window, 3 atlas generation iterations, temporal Gaussian kernel with constant 1 week sigma, and 0.7 mm output isotropic resolution settings.

[2] MIRTK library: https://github.com/BioMedIA/MIRTK.

2.3 Automated Ocular Biometry

The proposed pipeline for fetal ocular biometry is summarised in Fig. 2. We first localise the orbit using 3D U-Net [4], then we fit a 3D line through the orbit centroids and calculate the standard ocular measurements [21]. Furthermore, the step for automated detection of outliers provides quality control of the segmentations and measurements.

For the orbit segmentation module, the 3D U-Net [4] architecture consists of 5 encoding-decoding levels with 32, 64, 128, 256 and 512 channels, respectively. Each encoder block consists of 2 repeated blocks of $3 \times 3 \times 3$ convolutions (with a stride of 1), instance normalisation and LeakyReLU activations. The first two down-sampling blocks contains a $2 \times 2 \times 2$ average pooling layers, while the others use $2 \times 2 \times 2$ max pooling layers. The decoder blocks have a similar architecture as the encoder blocks, followed by upsampling layers. The model outputs an N-channel 3D image, corresponding to our 2 classes: background and fetal orbits. The network is implementated in PyTorch[3].

The orbit masks were created manually by a trained clinician for 20 cases. The network was trained on 19 3D reconstructed images and 1 case was used for validation. The training was performed for 100 epochs with TorchIO augmentation [17] including affine transformations ($\pm 180°$ rotations and $0.9 - 1.1$ scaling), bias field and motion artifacts ($< 5°$ rotations and < 2.0 translations).

Fig. 2. Proposed pipeline for automated ocular biometry.

Next, the output 3D segmentations are post-processed using morphological filtering. The two largest components with $\pm 35\%$ difference in volume are selected as the orbits (to account for a potential intensity variability due to the presence of a bias field). The calculation of ocular biometry is performed by fitting a line through the orbit centroids followed by detection of intersection points, calculation of line segment length and extraction of the standard 2D metrics: ocular diameter (OD), binocular distance (BOD), interocular distance (IOD), see Fig. 7.f. In addition, we also calculate volumes of the orbits (OV). Similarly to [21], the mean OD and OV values are computed as an average between the left and right orbits.

Outlier detection is based on three inclusion conditions: (i) the number of detected components should not exceed 5, which is an indicator of not well defined ocular features due to low image quality; (ii) the sizes of the right and left (R/L) orbits should be comparable within $\pm 15\%$ difference in terms of both volumes and OD values; (iii) the extracted metrics should be within $\pm 30\%$ window of the GA-specific curve values [21].

[3] PyTorch: https://pytorch.org.

In order to confirm the correlation between the intracranial and ocular volumetry, we trained a 3D U-Net with the same architecture for brain extraction. The training was performed for 400 epochs with augmentation in 3 steps using semi-supervised approach. At the first stage we used 60 fetal brain SVR images with manual segmentations of the intracranial volume available from other research projects. Next, the results of testing on the entire cohort (291) were examined and successful brain masks that included the entire intracranial volume were used in the next stage of training. All output intracranial brain masks for good quality reconstruction cases were visually inspected by a trained researcher and manually refined in 32 cases due to the presence of errors.

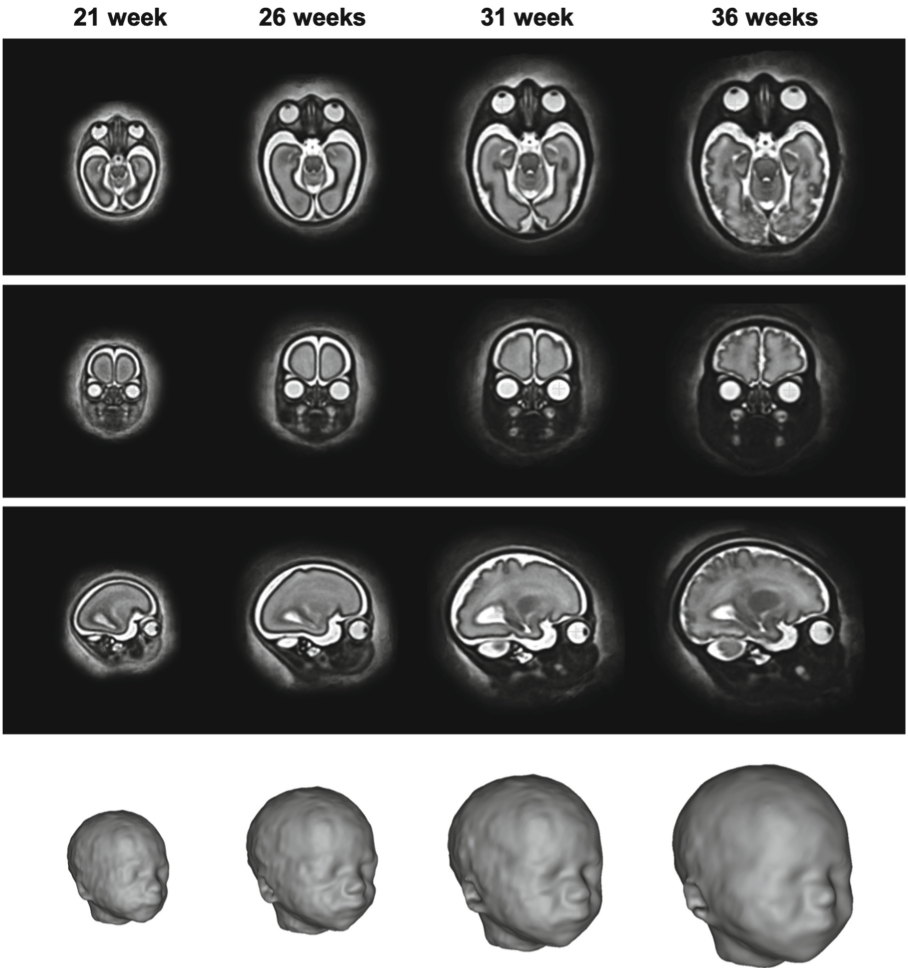

Fig. 3. The generated spatio-temporal atlas of the fetal craniofacial feature development at 21, 26, 31 and 36 weeks GA along with the corresponding face masks.

3 Results and Discussion

3.1 Spatio-Temporal Atlas of Fetal Craniofacial Feature Development

The generated spatio-temporal atlas of the head ROI at 21, 26, 31 and 36 weeks GA is shown in Fig. 3. The corresponding presented face masks were created semi-manually using combination of thresholding, manual refinement and label propagation from one to the rest of the GA timepoints. The atlas was inspected by two clinicians trained in fetal MRI who confirmed that all craniofacial features are correct, well defined and have high contrast. The atlas will be available online at the SVRTK data repository[4].

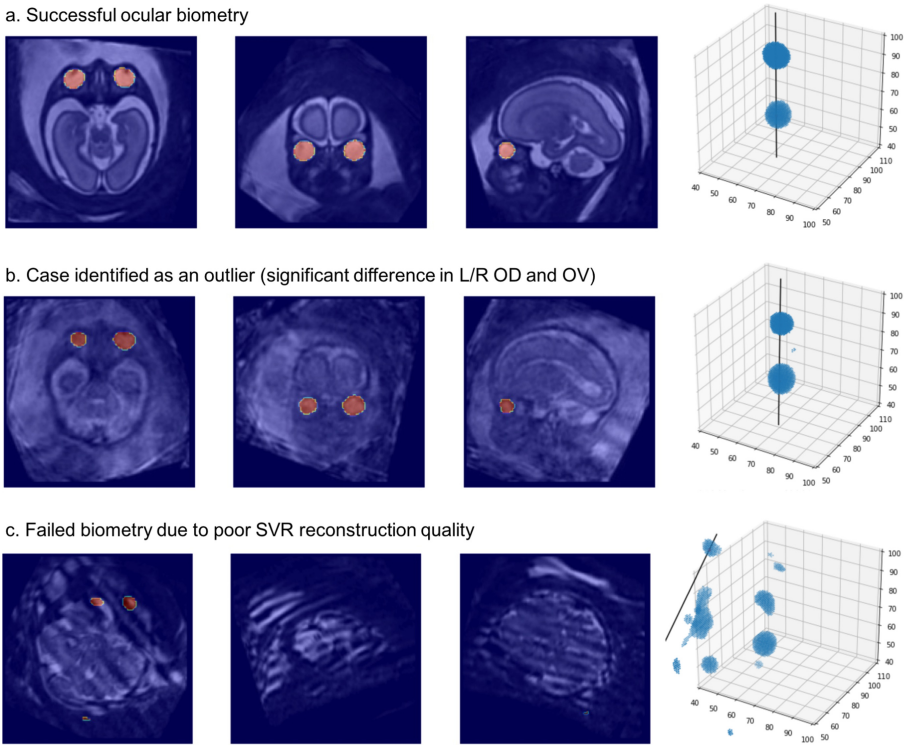

Fig. 4. Examples of a successful biometry output (a), a case detected as an outlier due to R/L OD and OV differences (b) and a completely failed case (c). The illustrations show the original 3D SVR reconstructions with orbit mask overlay and the corresponding 3D model with the fitted 3D line.

[4] SVRTK fetal and neonatal MRI data repository: https://gin.g-node.org/SVRTK.

3.2 Eye Biometry

Figure 4 shows an example of a successful (Fig. 4a) and failed (Fig. 4c) biometry output as well as a case that was automatically identified as an outlier (Fig. 4b) due to the difference between right and left OD and OV values. The completely failed case (Fig. 4c) was also automatically detected since there were > 5 components in the 3D U-Net output. This was caused by the low image quality due to the insufficient number of input stacks and the extreme motion that could not be resolved by SVR reconstruction [5].

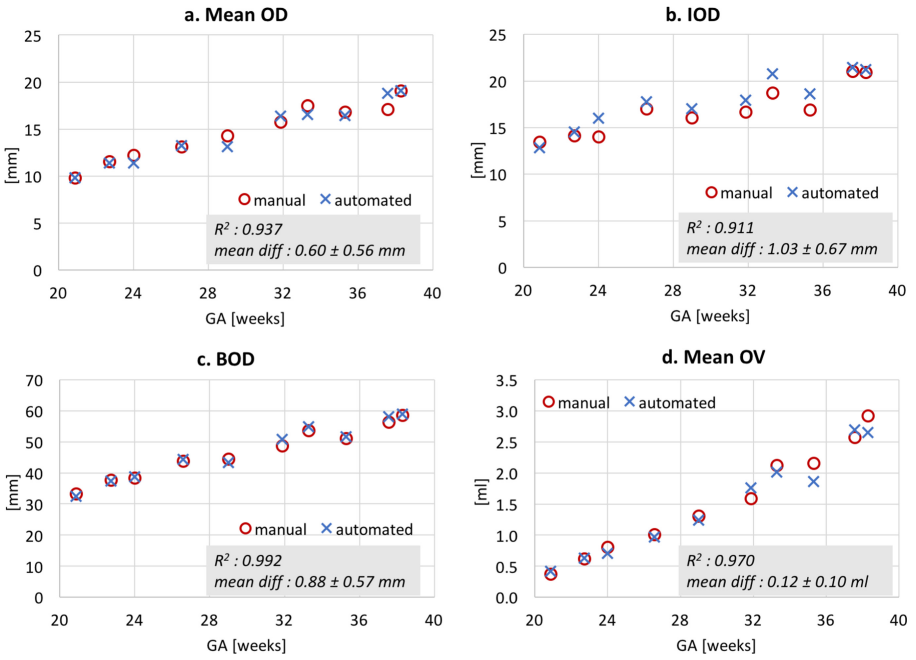

Fig. 5. Comparison between automated and manual measurements for mean OD (a), IOD (b), BOD (c) and mean OV (d) on 10 randomly selected cases.

The performance of the proposed pipeline of ocular biometry (Sect. 2.3) was evaluated on 10 randomly selected cases (quality grade group 3 to 4) from different GA groups, with automated biometry outputs compared to manual measurements. The corresponding results, presented in Fig. 5, show reasonably low absolute and relative differences (0.60 ± 0.56 mm and $3.99 \pm 3.49\%$ for mean OD, 1.03 ± 0.67 mm and $6.06 \pm 3.96\%$ for IOD and 0.88 ± 0.57 mm and $1.88 \pm 1.10\%$ for BOD, see Fig. 5) and high correlation ($R^2 > 0.91$) between the automated and manual measurements for all metrics. The slightly higher IOD and lower OV values in the automated output are primarily caused by more conservative automated segmentations that exclude the boundary around the orbits. It should be

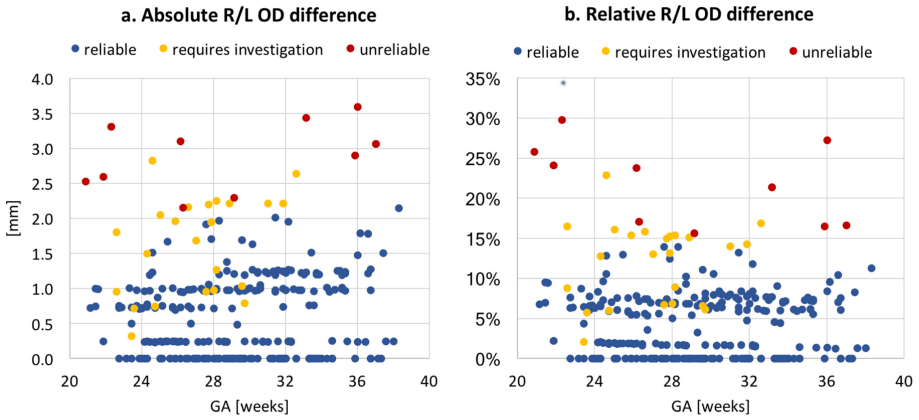

Fig. 6. Absolute (a) and relative (b) differences in right/left OD measurements. The cases in yellow and red are outliers with significant differences in either R/L OD or OV or both, respectively, while the cases in blue can be considered to be reliable for interpretation. (Color figure online)

noted that both manual and automated mean OD and OV measurements were characterised by R/L orbit differences.

The outlier detection step identified 4 cases where the reconstruction completely failed and 34 cases with high differences in right/left orbit metrics. An illustration of the absolute and relative R/L orbit OD differences for all cases is given in Fig. 6 with the outliers highlighted in yellow and red depending on whether there is a difference between either OD or OV or both these measurement. The discrete appearance of the difference values is related to the voxel size of the input images since the 2D distances are computed as voxels between intersection points along the fitted lines. The average quality scores (the manual grading in Fig. 1) in failed (4 cases), outlier (34 cases) and normal (253 cases) groups are 1.0 ± 0.0, 3.1 ± 0.9 and 3.6 ± 0.5, respectively. Notably, in addition to motion artefacts, the primary cause of the R/L differences was the presence of a strong bias field which was not taken into account during image quality grading.

3.3 Growth Charts

Prior to the analysis of growth trajectories, all automated eye segmentations and biometry results were also inspected manually which confirmed the similar size of the detected orbits in 253 cases. The automatically detected failed and outlier cases (Sect. 3.2) were excluded. The growth trajectories constructed from the selected 253 cases, shown in Fig. 7, include ocular biometry (mean OD, IOD, BOD), mean OV and total intracranial volume. The trajectories of all indices show high agreement with the existing formulas for ocular indices [21].

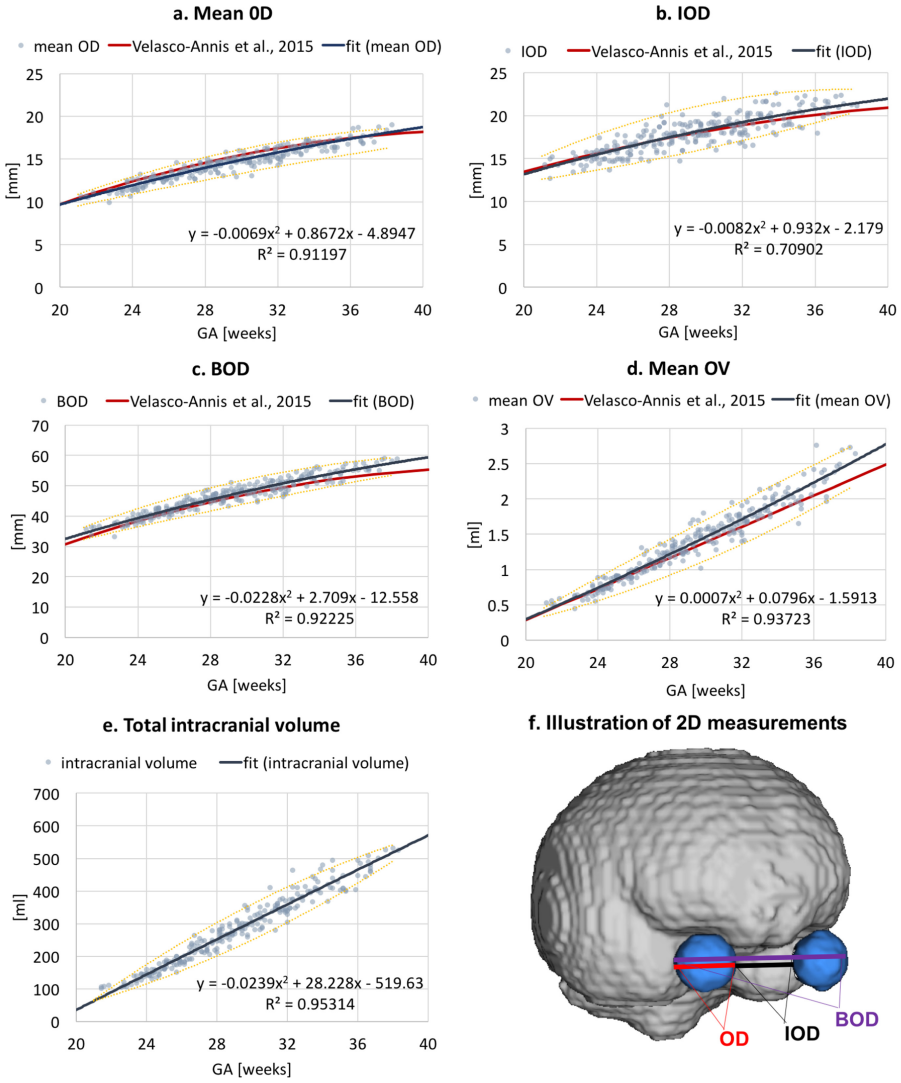

Fig. 7. Growth charts for 253 subjects from [20; 38] week GA range for mean OD (a), IOD (b), BOD (c), mean OV (d) and total intracranial volume (e) extracted from the automated measurements. The illustration of the measurements is given in (f).

4 Conclusions

In summary, we have presented the first spatio-temporal atlas of fetal cranio-facial feature development from 21 to 36 weeks GA which extends the existing brain-only fetal MRI atlases. The atlas will be available online at the SVRTK

data repository[5]. We also showed that fully automated 3D CNN-based ocular biometry can be used for processing large cohort datasets as an alternative to manual measurements. In addition, the proposed solution for detection of outliers provides the means to control interpretation of the outputs of automated processing by highlighting potentially unreliable results that require manual editing. The outlier cases with significant deviations that can occur due to either low image quality, failed segmentation or anomalies should be manually inspected, if required. The growth charts from the automatically derived ocular indices showed high correlation to the previously reported trends [21]. Our future work will focus on further automation of parcellation and biometry of craniofacial structures as well as analysis of abnormal cases.

Acknowledgments. We thank everyone who was involved in acquisition and analysis of the datasets at the Department of Perinatal Imaging and Health at King's College London. We thank all participating mothers.

This work was supported by the European Research Council under the European Union's Seventh Framework Programme [FP7/ 20072013]/ERC grant agreement no. 319456 dHCP project, the Wellcome/EPSRC Centre for Medical Engineering at King's College London [WT 203148/Z/16/Z)], the NIHR Clinical Research Facility (CRF) at Guy's and St Thomas' and by the National Institute for Health Research Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Ami, O., et al.: 3D magnetic resonance imaging of fetal head molding and brain shape changes during the second stage of labor. PLoS ONE **14**(5) (2019)
2. Arangio, P., et al.: Importance of fetal MRI in evaluation of craniofacial deformities. J. Craniofac. Surg. **24**(3), 773–776 (2013)
3. Avisdris, N., et al.: Automatic fetal ocular measurements in MRI. In: ISMRM 2021, p. 1190 (2021)
4. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
5. Cordero-Grande1, L., et al.: Automating motion compensation in 3T fetal brain imaging: localize, align and reconstruct. In: ISMRM 2019, p. 1000 (2019)
6. Ebner, M., et al.: An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain MRI. Neuroimage **206**(Oct.) (2020)
7. Ettema, A., et al.: Prenatal diagnosis of craniomaxillofacial malformations: a characterization of phenotypes in trisomies 13, 18, and 21 by ultrasound and pathology. Cleft Palate-Craniofac. J. **47**(2), 189–196 (2010)

[5] SVRTK fetal and neonatal MRI data repository: https://gin.g-node.org/SVRTK.

8. Gholipour, A., et al.: A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. Nat. Sci. Rep. **7**(476), 1–13 (2017)
9. Grigorescu, I., et al.: Harmonized segmentation of neonatal brain MRI. Front. Neurosci. **15**, 565 (2021)
10. Khalili, N., et al.: Automatic brain tissue segmentation in fetal MRI using convolutional neural networks. Magn. Reson. Imaging **64**, 77–89 (2019)
11. Kuklisova-Murgasova, M., et al.: Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. MedIAn **16**(8), 1550–1564 (2012)
12. Kul, S., et al.: Contribution of MRI to ultrasound in the diagnosis of fetal anomalies. J. Magn. Reson. Imaging **35**(4), 882–890 (2012)
13. Kyriakopoulou, V., et al.: Normative biometry of the fetal brain using magnetic resonance imaging. Brain Struct. Funct. **222**(5), 2295–2307 (2016). https://doi.org/10.1007/s00429-016-1342-6
14. Makropoulos, A., et al.: The dHCP: a minimal processing pipeline for neonatal cortical surface reconstruction. Neuroimage **173**, 88–112 (2018)
15. Mossey, P., Castilla, E.E.: Global registry and database on craniofacial anomalies Report of a WHO Registry Meeting on Craniofacial Anomalies Human Genetics Programme Management of Noncommunicable Diseases (2003)
16. Payette, K., Kottke, R., Jakab, A.: Efficient multi-class fetal brain segmentation in high resolution MRI reconstructions with noisy labels. In: Hu, Y., et al. (eds.) ASMUS/PIPPI -2020. LNCS, vol. 12437, pp. 295–304. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60334-2_29
17. Pérez-García, F., et al.: TorchIO: a Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. arXiv (March 2020)
18. Price, A., et al.: The developing Human Connectome Project (dHCP): fetal acquisition protocol. In: ISMRM 2019 (2019)
19. Robinson, A.J., et al.: MRI of the fetal eyes: morphologic and biometric assessment for abnormal development with ultrasonographic and clinicopathologic correlation. Pediatr. Radiol. **38**(9), 971–981 (2008)
20. Schuh, A., et al.: Unbiased construction of a temporally consistent morphological atlas of neonatal brain development. bioRxiv (2018)
21. Velasco-Annis, C., et al.: Normative biometrics for fetal ocular growth using volumetric MRI reconstruction. Prenat. Diagn. **35**(4), 400–408 (2015)
22. Wright, R., et al.: LSTM spatial co-transformer networks for registration of 3D fetal US and MR brain images. MICCAI **2018**, 107–116 (2018)