# Towards Realistic Ultrasound Fetal Brain Imaging Synthesis

**Michelle Iskandar**[*2] **Harvey Mannering**[†1] **Zhanxiang Sun**[‡1] **Jacqueline Matthew** [2] **Hamideh Kerdegari** [2] **Laura Peralta**[2] **Miguel Xochicale**[§1]
[1] *University College London* [2] *King's College London*

**Editors:** Under Review for MIDL 2023

## Abstract

Prenatal ultrasound imaging is the first-choice modality to assess fetal health. Medical image datasets for AI and ML methods must be diverse (i.e. diagnoses, diseases, pathologies, scanners, demographics, etc), however there are few public ultrasound fetal imaging datasets due to insufficient amounts of clinical data, patient privacy, rare occurrence of abnormalities in general practice, and limited experts for data collection and validation. To address such data scarcity, we proposed generative adversarial networks (GAN)-based models, diffusion-super-resolution-GAN and transformer-based-GAN, to synthesise images of fetal ultrasound brain planes from one public dataset. We reported that GAN-based methods can generate 256x256 pixel size of fetal ultrasound trans-cerebellum brain image plane with stable training losses, resulting in lower FID values for diffusion-super-resolution-GAN (average 7.04 and lower FID 5.09 at epoch 10) than the FID values of transformer-based-GAN (average 36.02 and lower 28.93 at epoch 60). The results of this work illustrate the potential of GAN-based methods to synthesise realistic high-resolution ultrasound images, leading to future work with other fetal brain planes, anatomies, devices and the need of a pool of experts to evaluate synthesised images. Code, data and other resources to reproduce this work are available at https://github.com/budai4medtech/midl2023.

**Keywords:** Medical Image Synthesis, Ultrasound Fetal Imaging, GANs

## 1. Introduction

Prenatal imaging is performed to assess various aspects of pregnancy, including confirmation of the pregnancy, screening for developmental defects, and investigation of pregnancy complications (Kline–Fath and Bitters, 2007). In the last decade, the fields of machine learning (ML) and artificial intelligence (AI) have been successful to model intelligent behaviors with minimal human interference (Hamet and Tremblay, 2017). Particularly, automatic classification of fetal ultrasound planes and fetal head biometric measurement (Burgos-Artizzu et al., 2020b; Sin, 2018; Fiorentino et al., 2022). Despite such advances, there are few challenges faced in prenatal imaging: (a) the accuracy of recorded measurements which can be caused by differences in intra-view variability of imaging equipment and inter-observer variability of sonographer skills (England, 2015; Sarris et al., 2012; Villar et al., 1989; Kesmodel, 2018), (b) availability of expert clinicians or trained technicians to select, to classify and to validate regions of interest (Burgos-Artizzu et al., 2020a), (c) the insufficient and limited amount of clinical data (Jang et al., 2018; Sin, 2018; He et al., 2021), (d) data accessibility

[*] Contributed equally
[†] Contributed equally
[‡] Contributed equally
[§] Contributed equally

© 2023 M. Iskandar, H. Mannering, Z. Sun, J.M. , H.K. , L. Peralta & M. Xochicale.

due to patient privacy or protection of personal health information (Shin et al., 2018), and (e) the cost of acquisition of clinical data as it requires expensive imaging equipment and experts for data collection and validation (Wang et al., 2019; Kim et al., 2019). Given the advances with generative adversarial networks (GAN) methods to handle problems in medical reconstructions, image resolution, enhancement, segmentation, lesion detection, data simulation or classification (AlAmir and AlGhamdi, 2022), we hypnotise that realistic ultrasound imaging can address challenges in data scarcity, accessibility and expensiveness. For instance, Eli et al. (2017) proposed a method of generating freehand ultrasound image simulation using a spatially conditioned GAN. Kazeminia et al. (2020) presented a review of the state-of-the-art research in GAN in medical imaging for classification, denoising, reconstruction, synthesis, registration, and detection. Montero et al. (2021) proposed a method to generate fetal brain US images using an unconditional GAN, StyleGAN2, specifically to improve the fine-grained plane classification, specifically the trans-thalamic and trans-ventricular plane. Hence, the aim of this work is to show the potential of GAN-based methods to generate realistic ultrasound fetal trans-cerebellum brain plane imaging with small datasets.

## 2. Methods and datasets

### 2.1. Diffusion-Super-Resolution-GAN (DSR-GAN)

We use a Denoising Diffusion Probabilistic Model (DDPM) (Ho et al., 2020) followed by a Super-Resolution-GAN (Ledig et al., 2017). To reduce computation time, we finetune a pretrained DDPM to produce 128x128 pixel images and then scale them up to 256x256 using SRGAN. After the DDPM and before SRGAN, histogram matching (Castleman, 1996) is applied to ensure that the colour distribution of the synthetic images matches the colour distribution of the real images.

### 2.2. Transformer-based-GAN (TB-GAN)

We use StyleSwin model, which features Swin Transformer layers designed to capture high quality details of the original images while simultaneously reducing memory usage, enabling synthesized images of higher resolutions (Zhang et al., 2022). Differentiable data augmentation (DiffAug) and adaptive pseudo augmentation (APA) are implemented to combat discriminator over-fitting due to limited data and ensure stability in training process (Zhao et al., 2020; Jiang et al., 2021).

### 2.3. Image Quality Assessment

Quality of synthesised images are evaluated with Fréchet inception distance (FID), measuring the distance between distributions of synthesised and original images (Heusel et al., 2017). The lower the FID number is, the more similar the synthesised images are to the original ones. FID metric showed to work well for fetal head ultrasound images compared to other metrics (Bautista et al., 2022).

### 2.4. Datasets

Trans-cerebellum brain plane ultrasound images from Voluson E6 were used for this work, consisting of 408 training images (Burgos-Artizzu et al., 2020a,b). Scans were collected by multiple operators of similar skill level at BCNatal during standard clinical practice between October 2018 and April 2019. DICOM images were collected and anominsied using png format, resulting in images of various pixel size (e.g., 692x480, 745x559, and 961x663).

## 3. Experiments: Design and results

Diffusion model was finetuned for 10000 epochs with the Adam optimiser to then train SRGAN for 200 epochs from scratch with the Adam optimiser. The images used to train both models are flipped horizontally, zoomed and rotated randomly to increase the variety of the dataset(Fig 1c). Transfer learning is used when training StyleSwin. The model was firstly pre-trained on Trans-thalamus plane for 500 epochs as contains larger number of images (1072). Then, the model was fine-tuned on Trans-cerebellum plane images for an additional 200 epochs. Adam optimizer was also used during both pre-training and fine-tuning stages, following the two time-scale update rule with learning rates of 1e-4 for the discriminator and 1e-5 for the generator (van den Heuvel et al., 2018).

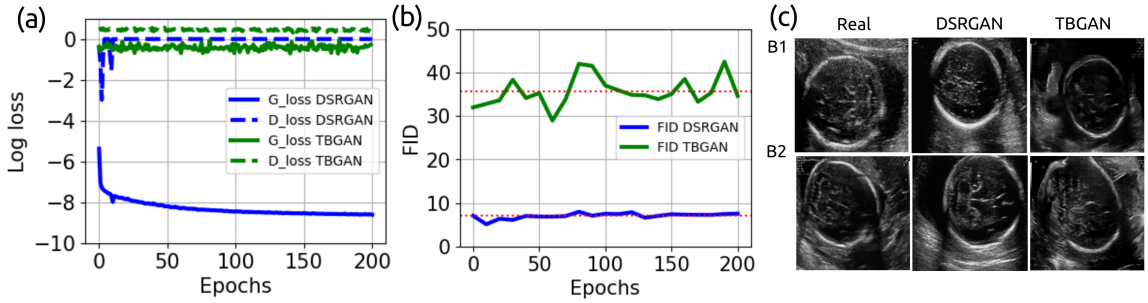


Figure 1: Results from Diffusion-Super-resolution-GAN (DSR-GAN) and transformer-based-GAN (TB-GAN): (a) Training losses for Generator and Discriminator networks, (b) FID scores, and (c) 256x256 pixel size trans-cerebellum images of two randomised batches (B1, B2) of real and synthesised (DSR-GAN and TB-GAN).

## 4. Conclusions and future work

Synthesising fetal brain images with the diffusion-Super-Resolution-GAN and transformer-based-GAN methods were successful, generating images of 256x256 pixel size resolution with stable loss values and resulting in lower FID values for diffusion-super-resolution-GAN (average 7.04 and lower FID 5.09 at epoch 10) compared to FID values of transformer-based-GAN (average 36.02 and lower 28.93 at epoch 60). Such results suggest, as future work, the potential to synthesise realistic higher-resolution fetal ultrasound images for other anatomies and ultrasound-devices.

## References

*Human-level Performance On Automatic Head Biometrics In Fetal Ultrasound Using Fully Convolutional Neural Networks*, 7 2018. IEEE. ISBN 978-1-5386-3646-6. doi: 10.1109/EMBC.2018.8512278.

Manal AlAmir and Manal AlGhamdi. The role of generative adversarial network in medical image analysis: An in-depth survey. *ACM Comput. Surv.*, mar 2022. ISSN 0360-0300. doi: 10.1145/3527849. URL https://doi.org/10.1145/3527849. Just Accepted.

Thea Bautista, Jacqueline Matthew, Hamideh Kerdegari, Laura Peralta Pereira, and Miguel Xochicale. Empirical study of quality image assessment for synthesis of fetal head ultrasound imaging with dcgans, 2022. URL https://arxiv.org/abs/2206.01731.

Xavier P. Burgos-Artizzu, David Coronado-Gutiérrez, Brenda Valenzuela-Alcaraz, Elisenda Bonet-Carne, Elisenda Eixarch, Fatima Crispi, and Eduard Gratacós. Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Scientific Reports*, 10(1):10200, Jun 2020a. ISSN 2045-2322. doi: 10.1038/s41598-020-67076-5. URL https://doi.org/10.1038/s41598-020-67076-5.

Xavier P. Burgos-Artizzu, David Coronado-Gutierrez, Brenda Valenzuela-Alcaraz, Elisenda Bonet-Carne, Elisenda Eixarch, Fatima Crispi, and Eduard Gratacós. Fetal planes db: Common maternal-fetal ultrasound images. *Scientific Reports*, 6 2020b. doi: 10.5281/ZENODO.3904280. URL https://zenodo.org/record/3904280.

Kenneth R Castleman. *Digital image processing*. Prentice Hall Press, 1996.

Eli, Lee Li-Lin, Xie Weidi, Barratt Dean C., Vercauteren Tom, Noble J Alison Hu Yipeng, and Gibson. Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks. *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment*, pages 105–115, 2017.

NHS England. *Fetal anomaly screening programme handbook*. NHS Digital, 7 2015.

Maria Chiara Fiorentino, Francesca Pia Villani, Mariachiara Di Cosmo, Emanuele Frontoni, and Sara Moccia. A review on deep-learning algorithms for fetal ultrasound-image analysis, 2022. URL https://arxiv.org/abs/2201.12260.

Pavel Hamet and Johanne Tremblay. Artificial intelligence in medicine. *Metabolism*, 69: S36–S40, 4 2017. ISSN 00260495. doi: 10.1016/j.metabol.2017.01.011.

Fujiao He, Yaqin Wang, Yun Xiu, Yixin Zhang, and Lizhu Chen. Artificial intelligence in prenatal ultrasound diagnosis. *Frontiers in Medicine*, 8, 2021. ISSN 2296-858X. doi: 10.3389/fmed.2021.729978. URL https://www.frontiersin.org/article/10.3389/fmed.2021.729978.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in Neural Information Processing*

*Systems*, 30, 2017. URL https://proceedings.neurips.cc/paper/2017/file/8a1d694707eb0fefe65871369074926d-Paper.pdf.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/4c5bcfec8584af0d967f1ab10179ca4b-Paper.pdf.

Jaeseong Jang, Yejin Park, Bukweon Kim, Sung Min Lee, Ja-Young Kwon, and Jin Keun Seo. Automatic estimation of fetal abdominal circumference from ultrasound images. *IEEE Journal of Biomedical and Health Informatics*, 22:1512–1520, 9 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2017.2776116.

Liming Jiang, Bo Dai, Wayne Wu, and Chen Change Loy. Deceive d: Adaptive pseudo augmentation for gan training with limited data, 2021.

Salome Kazeminia, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. Gans for medical image analysis. *Artificial Intelligence in Medicine*, 109:101938, 2020. ISSN 0933-3657. doi: https://doi.org/10.1016/j.artmed.2020.101938. URL https://www.sciencedirect.com/science/article/pii/S0933365719311510.

Ulrik S. Kesmodel. Information bias in epidemiological studies with a special focus on obstetrics and gynecology. *Acta Obstetricia et Gynecologica Scandinavica*, 97:417–423, 4 2018. ISSN 00016349. doi: 10.1111/aogs.13330.

Mingyu Kim, Jihye Yun, Yongwon Cho, Keewon Shin, Ryoungwoo Jang, Hyun jin Bae, and Namkug Kim. Deep learning in medical imaging. *Neurospine*, 16:657–668, 12 2019. ISSN 2586-6583. doi: 10.14245/ns.1938396.198.

Beth Kline–Fath and Constance Bitters. Prenatal imaging. *Newborn and Infant Nursing Reviews*, 7:197–204, 12 2007. ISSN 15273369. doi: 10.1053/j.nainr.2007.09.002.

Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

Alberto Montero, Elisenda Bonet-Carne, and Xavier Paolo Burgos-Artizzu. Generative adversarial networks to improve fetal brain fine-grained plane classification. *Sensors (Basel, Switzerland)*, 21, 11 2021. ISSN 1424-8220. doi: 10.3390/s21237975.

I. Sarris, C. Ioannou, P. Chamberlain, E. Ohuma, F. Roseman, L. Hoch, D. G. Altman, and A. T. Papageorghiou. Intra- and interobserver variability in fetal ultrasound measurements. *Ultrasound in Obstetrics and Gynecology*, 39:266–273, 3 2012. ISSN 09607692. doi: 10.1002/uog.10082.

Hoo-Chang Shin, Neil A Tenenholtz, Jameson K Rogers, Christopher G Schwarz, Matthew L Senjem, Jeffrey L Gunter, Katherine P Andriole, and Mark H Michalski. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. *ArXiv*, abs/1807.10225, 2018.

Thomas L. A. van den Heuvel, Dagmar de Bruijn, Chris L. de Korte, and Bram van Ginneken. Automated measurement of fetal head circumference using 2D ultrasound images. *PLOS ONE*, July 2018. doi: 10.5281/zenodo.1327317. URL https://doi.org/10.5281/zenodo.1327317.

J Villar, J Repke, L Markush, W Calvert, and G Rhoads. The measuring of blood pressure during pregnancy. *American journal of obstetrics and gynecology*, 161:1019–24, 10 1989. ISSN 0002-9378. doi: 10.1016/0002-9378(89)90777-1.

Ruoyao Wang, Zhenghan Fang, Jiaqi Gu, Yi Guo, Shicong Zhou, Yuanyuan Wang, Cai Chang, and Jinhua Yu. High-resolution image reconstruction for portable ultrasound imaging devices. *EURASIP Journal on Advances in Signal Processing*, 2019:56, 12 2019. ISSN 1687-6180. doi: 10.1186/s13634-019-0649-x.

Bowen Zhang, Shuyang Gu, Bo Zhang, Jianmin Bao, Dong Chen, Fang Wen, Yong Wang, and Baining Guo. Styleswin: Transformer-based gan for high-resolution image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11304–11314, June 2022.

Shengyu Zhao, Zhijian Liu, Ji Lin, Jun-Yan Zhu, and Song Han. Differentiable augmentation for data-efficient gan training, 2020.