Fusing Radiomic Features with Deep Representations for Gestational Age Estimation in Fetal Ultrasound Images

Fangyijie Wang[1,2]✉, Yuan Liang[1,4], Sourav Bhattacharjee[3], Abey Campbell[4], Kathleen M. Curran[1,2], and Guénolé Silvestre[1,4]

[1] Research Ireland Centre for Research Training in Machine Learning
fangyijie.wang@ucdconnect.ie
[2] School of Medicine, University College Dublin, Dublin, Ireland
[3] School of Veterinary Medicine, University College Dublin, Dublin, Ireland
[4] School of Computer Science, University College Dublin, Dublin, Ireland

Abstract. Accurate gestational age (GA) estimation, ideally through fetal ultrasound measurement, is a crucial aspect of providing excellent antenatal care. However, deriving GA from manual fetal biometric measurements depends on the operator and is time-consuming. Hence, automatic computer-assisted methods are demanded in clinical practice. In this paper, we present a novel feature fusion framework to estimate GA using fetal ultrasound images without any measurement information. We adopt a deep learning model to extract deep representations from ultrasound images. We extract radiomic features to reveal patterns and characteristics of fetal brain growth. To harness the interpretability of radiomics in medical imaging analysis, we estimate GA by fusing radiomic features and deep representations. Our framework estimates GA with a mean absolute error of 8.0 days across three trimesters, outperforming current machine learning-based methods at these gestational ages. Experimental results demonstrate the robustness of our framework across different populations in diverse geographical regions. Our code is publicly available on GitHub.

Keywords: Feature Fusion · Radiomics · Fetal Gestational Age · Ultrasound.

1 Introduction

Gestational age (GA) is the common term used during pregnancy to know the duration of gestation. This estimation is based on the assumption of an average fetal size at each stage of gestation. A normal pregnancy can range from 37 to 42 weeks [1]. Accurate dating of pregnancy is crucial for effective pregnancy management throughout all stages, from the first trimester to delivery [21]. It is essential for identifying premature labor and postdated deliveries [21]. Therefore, precision estimation of GA enables the accurate scheduling of antenatal care for women, guides obstetric management decisions, and supports the proper interpretation of fetal growth assessments [23].

However, in many settings, women attend their antenatal care late in pregnancy or even upon delivery, which presents challenges in accurately estimating GA [21]. Ultrasound imaging is widely used for screening and monitoring of pregnant women, as well as fetal growth [18,22]. In the screening examination, biometric measurements of the fetus, including the crown-rump length (CRL), head circumference (HC), abdominal circumference (AC), and femur length (FL), are commonly calculated to estimate the GA [12,23].

Recent advances in deep learning (DL) algorithms for fetal ultrasound image analysis have shown the significant potential of deep representations in the accuracy of diagnosis and biometric estimation [8,20]. Several studies have employed regression algorithms to estimate GA from ultrasound images [15,14]. These methodologies primarily rely on deep representations extracted by DL models, except [10], which provides an automated multimodal pipeline to estimate biometric parameters and GA at the same time. DL models exhibit exceptional performance in feature extraction; however, their lack of interpretability has restricted the development of clinical applications in estimating GA. Prior studies have shown that ultrasound images obtained during pregnancy can offer valuable insights into fetal brain maturation, assisting in estimating GA for screening purposes [19,26]. Recent research [27] demonstrates that radiomic features capture physics-related characteristics, making them intrinsically interpretable. However, there is a gap in the existing research as no study has yet explored combining radiomic features with DL techniques for fetal ultrasound imaging analysis. To our knowledge, this study represents the pioneering effort in creating an interpretable pathway that combines radiomic features and deep representations to estimate GA in fetal ultrasound images.

In this paper, we present a novel feature fusion framework that combines radiomic features and deep representations of the fetal head extracted by a convolutional neural network (CNN) to improve the predictive accuracy of GA. We evaluate the performance of our framework using two different fetal head ultrasound datasets and present quantitative results obtained from machine learning (ML) and DL models for comparison. The main contributions of our work: (i) we propose the first-ever feature fusion framework that combines fetal ultrasound images and radiomics to estimate GA; (ii) we present a plug-and-play cross-attention module for the effective fusion of radiomic features and deep representations; (iii) finally, comprehensive experiments demonstrate the effectiveness of our fusion module in estimating GA on ultrasound data.

2   Methodology

2.1   Overview

An overview of our proposed fusion framework for estimating GA with radiomic features and deep representations is illustrated in Fig. 1. The framework contains three core components: (1) a pre-trained DL model, ConvNeXt [17], is used for deep representation learning; (2) a Python pipeline extracts radiomic features; (3) a cross-attention module for feature fusion.
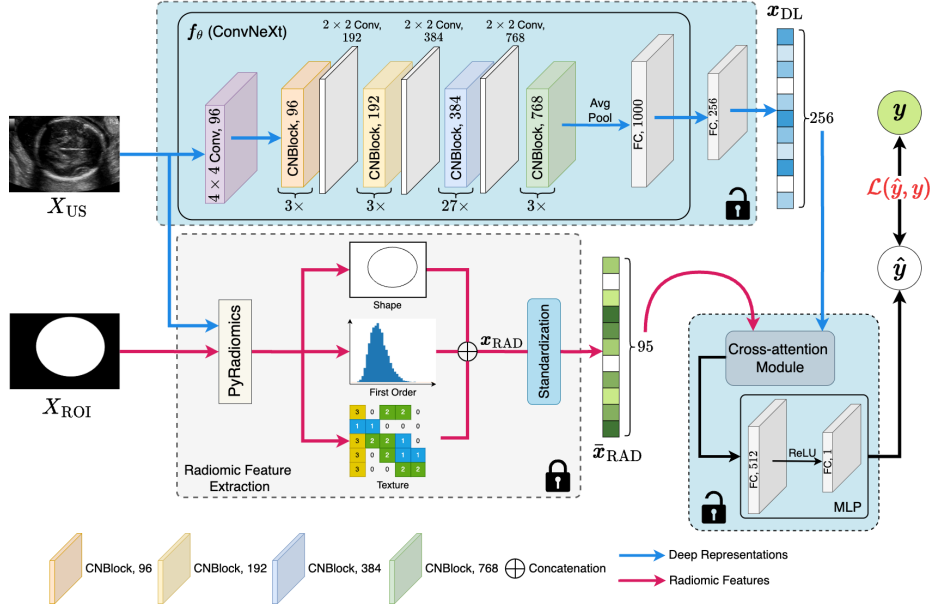
Fig. 1. Proposed fusion framework for GA estimation from fetal ultrasound images. **Blue Box**: Trainable parameters (DL model, cross-attention, MLP). **Gray Box**: Non-trainable parameters (Radiomics).

2.2 Deep Representation Learning

In our proposed framework, we employ the CNN topology to construct the DL model f_θ to learn deep representations that reflect high-level semantic information of ultrasound images, where θ represents the learnable parameters of the model. Specifically, we incorporate models such as ResNet18 [11], SwinTransformer [16], EfficientNet [24], MobileNet V3 [13], ViT [6], MaxViT [25], and ConvNeXt [17] for their proven effectiveness in image classification tasks. The DL model f_θ in our proposed framework adopts the ConvNeXt stack architecture with four layers. Each layer consists of a CNBlock followed by a convolution layer. The CNBlock layers have channel size of 96, 192, 384, and 768, respectively. The detailed illustration of the CNBlock structure is shown in Fig. 2(a).

When working with a grayscale fetal ultrasound image X_{US}, the initial step involves converting X_{US} into a 3-channel image denoted as $X_{US} \in \mathbb{R}^{3 \times n_h \times n_w}$, where n_h and n_w represent the height and width of the image, respectively. This conversion is essential to facilitate the fine-tuning of pre-trained DL models designed to process 3-channel input images. The input X_{US} is processed by the model f_θ, producing a deep representation vector that is subsequently fed to the ReLU activation and linearly combined to output $x_{DL} \in \mathbb{R}^{512}$. Subsequently, x_{DL} is fused with radiomic features $\bar{x}_{RAD}$ within the cross-attention module.
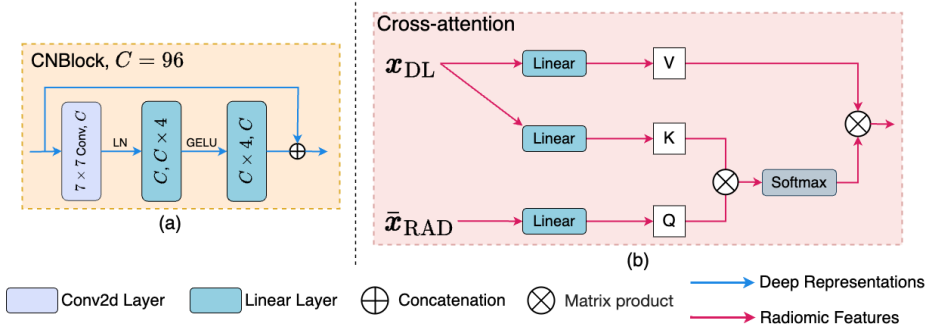
Fig. 2. Visualization of CNBlock and Cross-attention module. **(a)** is an example of CNBlock with channel size 96; **(b)** represents the cross-attention module for integrating radiomic features and deep representations. LN: Layer Normalization

2.3 Radiomic Features

Our dataset comprises fetal head ultrasound images and corresponding region of interest (ROI) masks, denoted as X_{US} and X_{ROI}, respectively. The ROI encompasses the full extent of the fetal head, as annotated by sonographers. The input X_{ROI} has dimensions of $1 \times n_{\mathrm{h}} \times n_{\mathrm{w}}$, where n_{h} and n_{w} correspond to the height and width of ROI image. The process of extracting radiomic features utilizes each pair of X_{US} and X_{ROI} as input. In this study, a total of 95 radiomic features $x_{\mathrm{RAD}} \in \mathbb{R}^{95}$, which capture surface-level details of the fetal head region, are extracted with the open-source tool, `pyradiomics`[5]. These extracted features are categorized into shape, statistical (first-order), and texture features [7,2,9].

The radiomic features are categorized into six groups: shape features (Shape2D), First-order features, texture features consisting of gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), and gray level dependence matrix (GLDM) features. Before the feature fusion step, we standardize radiomic features x_{RAD} by removing the mean and scaling to unit variance to obtain $\bar{x}_{\mathrm{RAD}} \in \mathbb{R}^{95}$. Subsequently, these standardized features $\bar{x}_{\mathrm{RAD}}$ are used as the input to the cross-attention module.

2.4 Feature Fusion

We pass features to a cross-attention module to perform fusion operation; see Fig. 2(b). The cross-attention module in the model has an embedding size of 512 and is linked to a multi-layer perceptron (MLP) module. The MLP module consists of two fully connected layers with a hidden feature size of 512 and the ReLU activation layer. Firstly, input feature $\bar{x}_{\mathrm{RAD}}$ is linearly projected to queries $q = W_{\mathrm{Q}}\bar{x}_{\mathrm{RAD}} + b_{\mathrm{Q}}$ ($q \in \mathbb{R}^{d_e}$), and feature x_{DL} is linearly projected to keys $k = W_{\mathrm{K}}x_{\mathrm{DL}} + b_{\mathrm{K}}$ ($k \in \mathbb{R}^{d_e}$), and values $v = W_{\mathrm{V}}x_{\mathrm{DL}} + b_{\mathrm{V}}$ ($v \in \mathbb{R}^{d_e}$). d_e

---

[5] https://pyradiomics.re adthedocs.io/en/latest/

represents the dimension of hidden embeddings. In our framework, d_e is set to 512. Next, the scaled dot-product attention is employed in the cross-attention module. Mathematically, it can be expressed as:

$$X_{\mathrm{XA}} = \mathrm{softmax}\left(\frac{\boldsymbol{q}\boldsymbol{k}^\top}{\sqrt{d_{\boldsymbol{k}}}}\right)\boldsymbol{v} = A_{\mathrm{XA}}\boldsymbol{v} \tag{1}$$

where $A_{\mathrm{XA}} \in \mathbb{R}^{d_q \times d_k}$ is the attention weight matrix, and $d_{\boldsymbol{k}}$ is the feature dimension of $\boldsymbol{k}$. The attention score ($\boldsymbol{q}\boldsymbol{k}^\top \in \mathbb{R}^{d_q \times d_k}$) is computed between each query and all keys. Normalization is done across the key dimension to obtain attention weights $A_{\mathrm{XA}} = \mathrm{softmax}(\cdot)$. The MLP module processes the cross-attention module's output to predict GA (denoted as $\hat{y} \in \mathbb{R}$) as follows:

$$\boldsymbol{y}_{\mathrm{XA}} = \mathrm{ReLU}(W_{\mathrm{XA}} \cdot \mathrm{vec}(X_{\mathrm{XA}}) + b_{\mathrm{XA}}) \tag{2}$$

$$\hat{y} = W_{\mathrm{MLP}} \cdot \boldsymbol{y}_{\mathrm{XA}} + b_{\mathrm{MLP}} \tag{3}$$

where $\mathrm{vec}(\cdot)$ is the flatten function, and $\boldsymbol{y}_{\mathrm{XA}} \in \mathbb{R}^{512}$ is a feature vector.

2.5   Model Training and Loss Functions

We employ a fine-tuning strategy for our framework, initializing the DL models with pre-trained weights from ImageNet [5]. Only the trainable components in our proposed framework are fine-tuned; see Fig. 1. We optimize the likelihood of the $\hat{y}$ regression output and therefore use the least squares for the overall training loss: $\mathcal{L} = \sum_{i=1}^{M} \left(y^{\langle i \rangle} - \hat{y}^{\langle i \rangle}\right)^2$, where $\hat{y}^{\langle i \rangle}$ is the model predicted GA for the ith training input and $y^{\langle i \rangle}$ is the computed GA as described in Section 3.1. M is the size of the training set. The model with the best performance is saved for testing based on their predictive accuracy during the model training phase.

Machine Learning. In our extensive analysis, we build traditional ML models for predicting GA using radiomic features exclusively. These models include Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), AdaBoost, Ridge and Gradient Boosting Regression (GBR). Additionally, we perform feature selection, including recursive feature elimination (REF) and least absolute shrinkage and selection operator (LASSO), to improve the performance of ML models.

3   Experiments and Results

3.1   Datasets

We utilize two public datasets in this study: the Spanish trans-thalamic dataset (**ES-TT**) originates from two centers in Spain [4]; the **HC18** dataset is sourced from a database in Netherlands [12]. All scans were acquired from healthy singleton pregnancies. HC18 data was collected using two ultrasound machines from a

single center, while ESTT data was obtained from six machines across two different centers. The variability of the fetal head across various pregnancy trimesters poses a challenge in estimating GA. We combine two datasets into one and use a 70:30 ratio to split the data into training and testing sets. A subset of 100 images from the test set is randomly selected for validation purposes.

ES-TT. The images in this dataset are captured when pregnant women are in their second and third trimesters undergoing routine examinations [4,3]. The training set comprises 1086 images from 704 patients, while the test set includes 466 images from 391 patients. Students, experienced physicians, and radiologists annotated the ROI of the fetal head in all images. A file detailing the pixel calibration, p_{size} (in mm), for each image is also provided.

HC18. This dataset includes fetuses without any growth abnormalities. The images are captured by experienced sonographers [12]. In each image, the sonographer draws an ellipse as ROI to best fit the circumference of the fetal head. We utilize 799 annotated images that are publicly accessible. The training set has 697 images from 603 patients, while the test set has 300 images from 279 patients. The pixel calibration, p_{size} (in mm), is provided for each image.

Formula Calculated GA. The ES-TT dataset provides the pixel size p_{size} for every ultrasound image in millimeters, while the HC18 dataset presents the HC values in millimeters corresponding to each image. To obtain the GA for all images, we calculate the number of pixels, denoted as p_{num}, along the edge of the respective ROI in the input image X_{ROI}. Subsequently, we compute the corresponding HC in millimeters (mm) using the formula: $HC = p_{\text{num}} \times p_{\text{size}}$. We obtain the GA $(y^{\langle i \rangle})$ using the equation derived from previous research [21]:
$GA = \exp\left[0.05970 \times (\log_e(\text{HC}))^2 + 0.000000006409 \times (\text{HC})^3 + 3.3258\right].$

3.2   Implementation Details

The input image X_{US} and mask X_{ROI} are resized to the same resolution, 256×256. To decrease computational demands, X_{US} is normalized following the same process from [11]. The Adam optimizer is utilized with an initial learning rate of 10^{-5} following a weight decay of 10^{-6}. The batch size is fixed at 8, and the epoch is set to 60. Our framework is implemented using Pytorch library and trained on a single RTX 4090 GPU with 24 GB memory.

Data Augmentation. During model training, we apply common data augmentation techniques on the images within the training set. Specifically, these augmentation techniques randomly rotate the images within range $(-15°, 15°)$ and randomly flip them horizontally with a probability of 50%.

3.3   Results

To assess the accuracy of predicting GA, we use Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) as the evaluation metrics. We find that the correlation coefficient between MAE and RMSE is larger than 0.998. Therefore, we only report the MAE in this paper. We evaluate our method by comparing

Table 1. Test results of GA estimation using DL models with concatenation and cross-attention mechanisms. C: Concatenation. XA: Cross-attention.

DL Model	C	XA	MAE ↓	P-value
ResNet18 [15]	✗	✗	9.9±0.02	
EfficientNet	✗	✗	9.3±0.00	
MaxViT	✗	✗	10.7±0.02	$<10^{-3}$
SwinTrans	✗	✗	9.4±0.01	
ConvNeXt	✗	✗	8.6±0.01	
ResNet18	✓	✗	10.2±0.01	
EfficientNet	✓	✗	8.6±0.02	
MaxViT	✓	✗	9.6±0.04	$<10^{-3}$
SwinTrans	✓	✗	8.6±0.01	
ConvNeXt	✓	✗	8.7±0.01	
ResNet18	✗	✓	10.4±0.01	
EfficientNet	✗	✓	8.6±0.00	
MaxViT	✗	✓	8.7±0.00	−
SwinTrans	✗	✓	8.3±0.00	
ConvNeXt	✗	✓	**8.0±0.00**	

Table 2. Test results of GA estimation using ML models by adopting LASSO and RFE techniques to select radiomic features. SVM: Support Vector Machine. KNN: K-Nearest Neighbors. RF: Random Forest. GBR: Gradient Boosting Regression.

ML Model	MAE ↓	R^2 ↑
SVM (LASSO)	25.2±24.3	0.62
KNN (LASSO)	25.2±25.1	0.61
AdaBoost (LASSO)	26.8±23.6	0.60
Ridge (LASSO)	25.7±22.3	0.64
RF (LASSO)	21.7±23.0	0.69
GBR (LASSO)	21.8±22.9	0.69
SVM (RFE)	24.9±24.4	0.62
KNN (RFE)	24.5±24.6	0.62
AdaBoost (RFE)	26.5±23.7	0.61
Ridge (RFE)	25.8±22.3	0.64
RF (RFE)	21.4±23.1	0.69
GBR (RFE)	21.3±22.6	0.70

it with Baseline utilizing only images and Concatenation utilizing concatenated image and radiomic features.

From Table 1, we observe several key findings. Our proposed framework has achieved the best result, with an average MAE of 8.0 (days). Quantitative analysis shows a significant improvement in accurate GA estimation with the cross-attention module compared to both the concatenation module ($p < 0.001$) and the baseline model ($p < 0.001$). Our method enhances the baseline method, ResNet18 [15], by 1.9 days in accurately estimating GA. Despite ResNet18 and EfficientNet models exhibiting higher MAE values when utilizing the cross-attention module than the concatenation module, the overall accuracy across all architectures is significantly improved with the former ($p < 0.001$), showcasing the superiority of the cross-attention mechanism. It is notable that ConvNeXt, an efficient model with 3.7M parameters, demonstrates superior performance compared to four other DL-based feature extractors, as evidenced in Table 1.

Table 2 shows the test results of various ML algorithms without using the radiomic features. The statistical metric R-squared (R^2), also known as the coefficient of determination, is employed to assess the adequacy of a model in fitting the data. The R^2 value suggests that GBR (RFE) is a more effective model for capturing the radiomic features than other ML algorithms. However, all ML algorithms exhibit significantly higher MAE values in GA estimation compared to our proposed method, highlighting the superiority of our method over traditional ML algorithms.

For better understanding, we visualize the cross-attention maps of three test cases from the 1st, 2nd, and 3rd trimesters in Fig. 3. The bar charts are employed
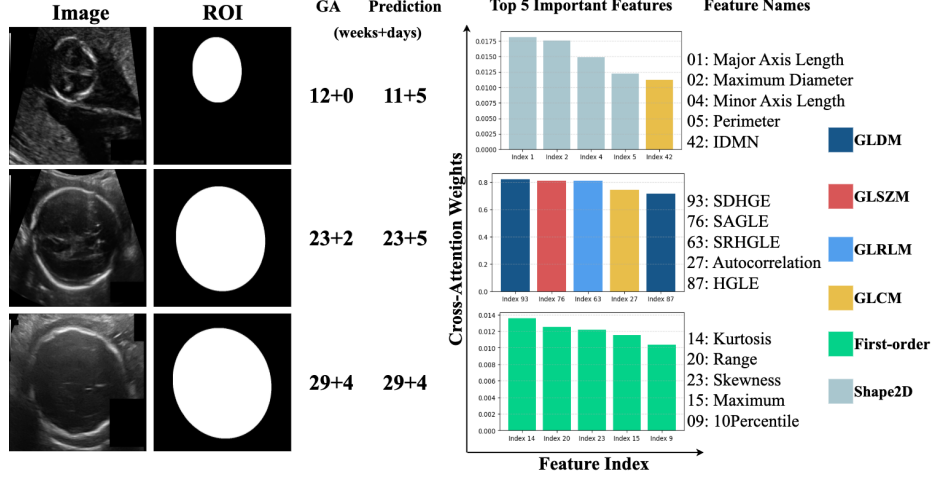
Fig. 3. The visualization of the interpretability of our method. IDMN: Inverse Difference Moment Normalized. SDHGE: Small Dependence High Gray Level Emphasis. SAGLE: Small Area High Gray Level Emphasis. SRHGLE: Short Run High Gray Level Emphasis. HGLE: High Gray Level Emphasis.

Table 3. Ablation study on the different feature fusion modules and pre-trained weights. C: Concatenation. XA: Cross-attention.

Model	Image	Pre-trained	C	XA	MAE ↓
ConvNeXt	✓	✗	✗	✗	28.9±0.04
ConvNeXt	✓	✓	✗	✗	8.6±0.01
ConvNeXt + C	✓	✓	✓	✗	8.7±0.01
ConvNeXt + XA (Ours)	✓	✓	✗	✓	**8.0±0.00**

to visualize the five most influential radiomic features for GA estimation, as determined by cross-attention weights. In the 1st and 3rd samples, Shape2D and First-order features are the primary contributors to the estimation. In contrast, in the 2nd sample, a diverse set of features such as GLDM, GLSZM, GLRLM, and GLCM play a significant role in the prediction.

Ablation Study. We investigate the effectiveness of three primary components in our framework: fine-tuning strategy, feature fusion strategy, and cross-attention module. We first set up a baseline framework without using these components. Table 3 illustrates the significance of fine-tuning the framework with a pre-trained DL model in enhancing the accuracy of predicting GA. The feature fusion strategy using the concatenation method demonstrates minimal significance in our framework. However, integrating cross-attention into our framework substantially enhances the accuracy of predicting GA. This indicates that the cross-attention module facilitates a more effective fusion of deep representations and radiomic features.

4   Conclusion

This study introduces a novel feature fusion framework that combines image and radiomic features to estimate fetal GA from ultrasound images. This framework can be developed into an interpretable tool for clinical use. We validate our framework on two fetal head ultrasound image datasets and achieve better results than those achieved by existing image-based DL and radiomic-based ML methods. Moreover, our framework can readily be utilized to analyze additional fetal anatomical structures in ultrasound, such as the fetal abdomen and femur. Hence, our proposed framework is a valuable algorithm that collaborates with a segmentation model to estimate GA, thereby benefiting patients and clinicians in clinical practice. Our future work includes enhancing the performance of GA estimations by utilizing a combination of multiple metrics obtained from various standard ultrasound planes across diverse cohorts. These experiments will be validated using larger datasets.

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

References

1. Gestational age. `https://medlineplus.gov/ency/article/002367.htm`, last accessed 2025-01-24
2. Aerts, H.J.W.L., Velazquez, E.R., Leijenaar, R.T.H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M.M., Leemans, C.R., Dekker, A., Quackenbush, J., Gillies, R.J., Lambin, P.: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications **5**(1), 4006 (Jun 2014)
3. Alzubaidi, M., Agus, M., Makhlouf, M., Anver, F., Alyafei, K., Househ, M.: Large-scale annotation dataset for fetal head biometry in ultrasound images. Data in Brief **51**, 109708 (Dec 2023)
4. Burgos-Artizzu, X.P., Coronado-Gutiérrez, D., Valenzuela-Alcaraz, B., Bonet-Carne, E., Eixarch, E., Crispi, F., Gratacós, E.: Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. Scientific Reports **10**(1), 10200 (Dec 2020)
5. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: proceedings of Conference on Computer Vision and Pattern Recognition (CVPR) (2009)
6. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. ICLR (2021)

7. Fan, M., Zhang, Y., Fu, Z., Xu, M., Wang, S., Xie, S., Gao, X., Wang, Y., Li, L.: A deep matrix completion method for imputing missing histological data in breast cancer by integrating dce-mri radiomics. Medical Physics **48**(12), 7685–7697 (2021)

8. Fiorentino, M.C., Villani, F.P., Di Cosmo, M., Frontoni, E., Moccia, S.: A review on deep-learning algorithms for fetal ultrasound-image analysis. Medical Image Analysis **83**, 102629 (Jan 2023)

9. van Griethuysen, J.J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R.G., Fillion-Robin, J.C., Pieper, S., Aerts, H.J.: Computational radiomics system to decode the radiographic phenotype. Cancer Research **77**(21), e104–e107 (Oct 2017)

10. Guo, X., Men, Q., Noble, J.A.: MMSummary: Multimodal Summary Generation for Fetal Ultrasound Video . In: proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI). vol. LNCS 15004. Springer Nature Switzerland (Oct 2024)

11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: proceedings of Conference on Computer Vision and Pattern Recognition (CVPR) (2016)

12. van den Heuvel, T.L.A., de Bruijn, D., de Korte, C.L., Ginneken, B.v.: Automated measurement of fetal head circumference using 2d ultrasound images. PLOS ONE **13**(8), 1–20 (Aug 2018)

13. Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L.C., Tan, M., Chu, G., Vasudevan, V., Zhu, Y., Pang, R., Adam, H., Le, Q.: Searching for mobilenetv3. In: proceedings of International Conference on Computer Vision (ICCV) (2019)

14. Lee, C., Willis, A., Chen, C., Sieniek, M., Watters, A., Stetson, B., Uddin, A., Wong, J., Pilgrim, R., Chou, K., Tse, D., Shetty, S., Gomes, R.G.: Development of a machine learning model for sonographic assessment of gestational age. JAMA Network Open **6**(1), e2248685–e2248685 (Jan 2023)

15. Lee, L.H., Bradburn, E., Craik, R., Yaqub, M., Norris, S.A., Ismail, L.C., Ohuma, E.O., Barros, F.C., Lambert, A., Carvalho, M., Jaffer, Y.A., Gravett, M., Purwar, M., Wu, Q., Bertino, E., Munim, S., Min, A.M., Bhutta, Z., Villar, J., Kennedy, S.H., Noble, J.A., Papageorghiou, A.T.: Machine learning for accurate estimation of fetal gestational age based on ultrasound images. npj Digital Medicine **6**(1), 36 (Mar 2023)

16. Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., Wei, F., Guo, B.: Swin transformer v2: Scaling up capacity and resolution. In: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). pp. 12009–12019 (Jun 2022)

17. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: proceedings of Conference on Computer Vision and Pattern Recognition (CVPR) (2022)

18. Loughna, P., Chitty, L., Evans, T., Chudleigh, T.: Fetal size and dating: Charts recommended for clinical obstetric practice. Ultrasound **17**(3), 160–166 (2009)

19. Namburete, A.I.L., Xie, W., Noble, J.A.: Robust regression of brain maturation from 3Dfetal neurosonography using CRNs. In: Fetal, Infant and Ophthalmic Medical Image Analysis. pp. 73–80. Springer International Publishing (2017)

20. Naz, S., Noorani, S., Jaffar Zaidi, S.A., Rahman, A.R., Sattar, S., Das, J.K., Hoodbhoy, Z.: Use of artificial intelligence for gestational age estimation: a systematic review and meta-analysis. Frontiers in Global Women's Health **6** (2025)

21. Papageorghiou, A.T., Kemp, B., Stones, W., Ohuma, E.O., Kennedy, S.H., Purwar, M., Salomon, L.J., Altman, D.G., Noble, J.A., Bertino, E., Gravett, M.G.,

Pang, R., Cheikh Ismail, L., Barros, F.C., Lambert, A., Jaffer, Y.A., Victora, C.G., Bhutta, Z.A., Villar, J., for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) : Ultrasound-based gestational-age estimation in late pregnancy. Ultrasound in Obstetrics & Gynecology **48**(6), 719–726 (2016)

22. Salomon, L.J., Alfirevic, Z., Berghella, V., Bilardo, C., Hernandez-Andrade, E., Johnsen, S.L., Kalache, K., Leung, K.Y., Malinger, G., Munoz, H., Prefumo, F., Toi, A., Lee, W., on behalf of the ISUOG Clinical Standards Committee: Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. Ultrasound in Obstetrics & Gynecology **37**(1), 116–126 (2011)

23. Sinha, P., Gupta, M., Sharma, R., Srivastava, K.R.: Comparison of estimation of gestational age by transverse cerebellar diameter with biparietal diameter in third trimester of pregnancy. Journal of South Asian Federation of Obstetrics and Gynaecology **12**(4), 235–238 (2020)

24. Tan, M., Le, Q.: Efficientnetv2: Smaller models and faster training. In: Meila, M., Zhang, T. (eds.) Proceedings of the 38th International Conference on Machine Learning. vol. 139, pp. 10096–10106. PMLR (Jul 2021)

25. Tu, Z., Talebi, H., Zhang, H., Yang, F., Milanfar, P., Bovik, A., Li, Y.: Maxvit: Multi-axis vision transformer. In: European Conference on Computer Vision (ECCV). pp. 459–479. Springer (2022)

26. Welp, A., Gembicki, M., Rody, A., Weichert, J.: Validation of a semiautomated volumetric approach for fetal neurosonography using 5DCNS+ in clinical data from > 1100 consecutive pregnancies. Child's Nervous System **36**(12), 2989–2995 (Dec 2020)

27. Zilka, T., Benesova, W.: Radiomics of pituitary adenoma using computer vision: a review. Medical & Biological Engineering & Computing **62**(12), 3581–3597 (Dec 2024)