DTU Library

Leveraging Shape and Spatial Information for Spontaneous Preterm Birth Prediction

Pegios, Paraskevas; Sejer, Emilie Pi Fogtmann; Lin, Manxi; Bashir, Zahra; Svendsen, Morten Bo Søndergaard; Nielsen, Mads; Petersen, Eike; Christensen, Anders Nymark; Tolsgaard, Martin; Feragen, Aasa

Published in:
Proceedings of The 4th International Workshop of Advances in Simplifying Medical Ultrasound (ASMUS)

Link to article, DOI:
10.1007/978-3-031-44521-7_6

Publication date:
2023

Document Version
Peer reviewed version

Leveraging Shape and Spatial Information for Spontaneous Preterm Birth Prediction

Paraskevas Pegios[1,5](✉), Emilie Pi Fogtmann Sejer[2], Manxi Lin[1], Zahra Bashir[3], Morten Bo Søndergaard Svendsen[2,4], Mads Nielsen[4,5], Eike Petersen[1], Anders Nymark Christensen[1], Martin Tolsgaard[2], and Aasa Feragen[1,5]

[1] Technical University of Denmark, Kongens Lyngby, Denmark
{ppar,afhar}@dtu.dk
[2] Region Hovedstaden Hospital, Copenhagen, Denmark
[3] Slagelse Hospital, Copenhagen, Denmark
[4] University of Copenhagen, Copenhagen, Denmark
[5] Pioneer Centre for AI, Copenhagen, Denmark

Abstract. Spontaneous preterm birth prediction from transvaginal ultrasound images is a challenging task of profound interest in gynecological obstetrics. Existing works are often validated on small datasets and may lack validation of model calibration and interpretation. In this paper, we present a comprehensive study of methods for predicting preterm birth from transvaginal ultrasound using a large clinical dataset. We propose a shape- and spatially-aware network that leverages segmentation predictions and pixel spacing information as additional input to enhance predictions. Our model demonstrates competitive performance on our benchmark, providing additional interpretation and achieving the highest performance across both clinical and machine learning baselines. Through our evaluation, we provide additional insights which we hope may lead to more accurate predictions of preterm births going forwards. Our source code is available at github.com/ppegiosk/SA-SonoNet-sPTB.

Keywords: Spontaneous Preterm Birth · Transvaginal Ultrasound · Transparency

1 Introduction

Spontaneous preterm birth (sPTB), usually defined as birth occurring before 37 weeks of gestation, is considered a pressing challenge, with substantial health, societal, and financial implications. Affecting millions of cases annually it is the key factor causing neonatal morbidity [24], as premature infants are vulnerable to several complications. These risks often necessitate prolonged hospitalization in neonatal intensive care units, with potentially adverse outcomes [13]. The ability to accurately predict sPTB is of paramount importance in the prevention of neonatal mortality and morbidity. By identifying pregnancies at risk, healthcare professionals can provide support to the affected infants and their families. Despite considerable efforts, predicting sPTB remains an open problem.

Cervical length (CL) measurements obtained from cervical ultrasound images currently serve as the clinical gold standard for sPTB prediction, with a threshold usually of CL<25 mm indicating an increased risk [7]. The success of machine learning opened up new opportunities [2, 26] by analyzing ultrasound images, electronic health data, and electrohysterogram signals, and emerging imaging modalities [20]. With the rise of deep learning, U-Net-based methods [25, 27] have demonstrated state-of-the-art results using ultrasound images. Yet, existing methods may suffer from potential risk of bias [28] due to their small effective sample size, and their lack of model transparency, and calibration evaluation for assessing the reliability of individual confidences.

We present a comprehensive analysis of methods that seek to predict sPTB using a class-balanced dataset of 7862 transvaginal ultrasound images. We argue that the shape, spatial, and textural information of the cervix are strong candidates for predicting sPTB. To this end, we leverage DTU-Net [16] for measuring CL, utilizing its ability to segment curvilinear structures. From this, we build a shape- and spatially-aware SA-SonoNet classifier, based on the SonoNet architecture [4], which uses predicted segmentations to enhance predictions following [15]. In summary, we contribute 1) an extensive evaluation of both clinical and machine learning baselines on a large clinical dataset and 2) a new method that outperforms existing approaches while providing additional explainability.

2 Related Work

Sonographic assessment of CL is considered the most accessible and accurate predictor for sPTB [7]. Nevertheless, improving population-wide screening for low-risk cases requires addressing the challenges of low sensitivity and prevalence of CL [23]. Thus, measurements of the uterocervical angle (UCA), i.e., the angle between the uterine wall and the cervical canal, have been explored [9, 10] or combined with CL [17, 21] as an additional biomarker, along with wall thickness measurements of the upper and lower uterine segment [1]. However, these approaches rely on the skills of the sonographer to measure biomarkers or require specialized protocols for data collection. We provide a) a method to automate CL measurements, and b) a second method to automatically predict sPTB directly from transvaginal ultrasound, simultaneously improving performance, speeding up the data collection, and reducing the demand on the clinician's competences.

Another branch of work [3, 6, 11] focuses on the analysis of cervical texture. In clinically hypothesized regions of interest (ROIs) hand-crafted textural features are extracted and used in conjunction with standard machine learning techniques. While these have shown promise, especially combined with CL [5], they are often hard to reproduce [2], and evaluated on small and highly class-unbalanced datasets with very few cases, and of higher image quality than the clinical standard. In this study, we implement and integrate a texture-based baseline into our analysis, evaluated on a large, class-balanced clinical dataset.

Recently, deep learning methods have been used to recognize cervical anatomy and interrogate sPTB predictions. In [27], a segmentation model predicts a bi-

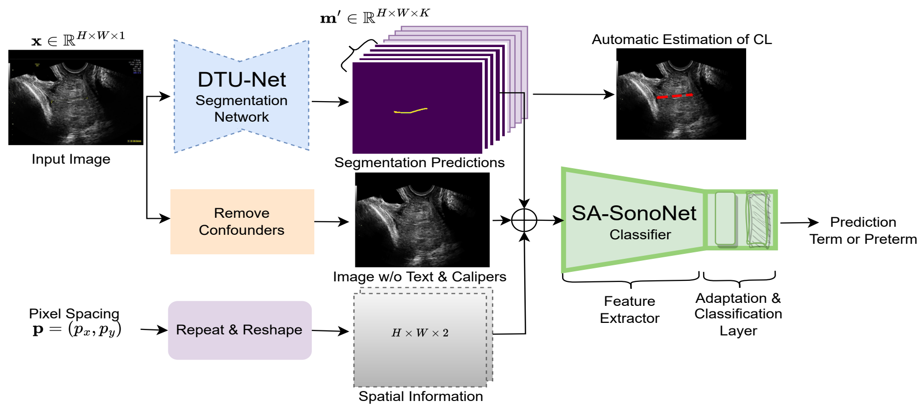


Fig. 1: Overview of the proposed method.

nary mask around the cervical canal, and then CL and UCA are estimated and serve as inputs for standard statistical methods. As an alternative to the binary mask, [8, 29] use multiple cervical coarse ROIs in order to provide feedback to clinicians. In [25], a multi-task framework is used both for classification and segmentation of the ROI surrounding the cervical canal. Although existing methods provide valuable information to clinicians, their utility in predicting sPTB is yet to be explored. In this work, we employ a DTU-Net [16] for recognizing cervical structures and leverage its predictions as additional inputs for our classification model, while we further inject it using spatial information.

3 Method

We summarize the architecture of our proposed method in Fig. 1. First, we use a DTU-Net [16] to automatically measure CL by detecting curvilinear structures in cervix images. Next, we introduce SA-SonoNet, a shape- and spatially-aware SonoNet [4] for predicting sPTB, which enhances performance by concatenating input images, segmentation predictions, and pixel spacing information.

3.1 Estimating Cervical Length from Curvilinear Segmentation

We quantify CL using a DTU-Net [16] that segments the curvilinear structures of the cervical canal (CC), inner boundary (IB) and outer boundary (OB), and the volumetric structure of the bladder (BL), from background. From the CC segmentation, CL is measured via its left and right-most points (see Fig. 1). Robust binary masks covering the area near the CC, as utilized in [25, 27], can be created by applying morphological dilation to the CC predictions (see Fig. 2).

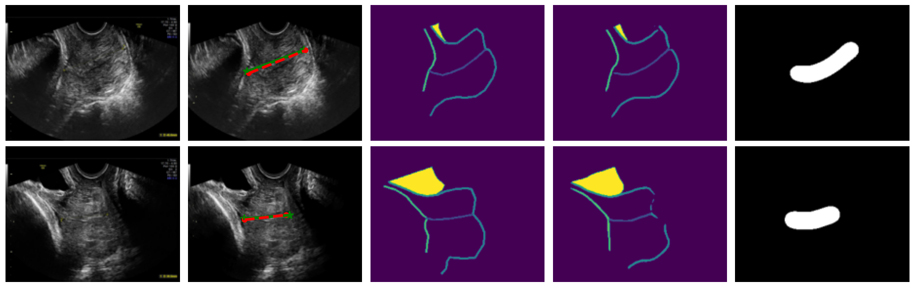


Fig. 2: Examples of a term (top) and a preterm (bottom) birth. **From left to right:** image, CL measurements (green: expert, red: predicteed), expert segmentations, predicted segmentations, generated binary masks. Binary masks are created by applying morphological dilation with a radius of 4 mm.

3.2 Injecting Shape and Spatial Information Into SonoNet

We hypothesize that the shape, size, and texture of the cervix are predictive of sPTB. Recognizing the inherent texture bias [14] of convolutional neural networks (CNNs), we inject shape and spatial information into SonoNet [4], a state-of-the-art CNN for ultrasound standard plane classification. As classification of standard plane quality benefits from combining images with predicted segmentations in [15], we similarly include the cervical shape information. As sonographers adjust image resolution during examinations, and we expect cervix size and texture to depend on resolution, we include pixel spacing as an input feature.

Let $\mathbf{x} \in \mathbb{R}^{H\times W\times 1}$ be a grayscale image with height H and width W, with pixel spacing information $\mathbf{p} = (p_x, p_y)$ representing the physical distance between the centers of each 2D pixel. We assume that we have access to a trained segmentation network g – in our experiments, the DTU-Net discussed in Section 3.1. During inference, g predicts a segmentation map $g(\mathbf{x}) = \mathbf{m} \in \mathbb{R}^{H\times W\times L}$, where L is the number of segmentation labels and $\mathbf{m}_{x,y}$ represents the probability distribution for the pixel at position (x, y) across the set of learned segmentation labels. We keep only the K segmentation predictions, that are relevant to the classification task, i.e, $\mathbf{m}' \in \mathbb{R}^{H\times W\times K}$. In our experiments, $K = 5$. Our goal is to learn a mapping $f(\mathbf{x}, \mathbf{m}', \mathbf{p}) \mapsto y$ where y indicates a predicted target and f is a SonoNet classifier. In practice, the pixel-spacing values (p_x, p_y) are repeated and reshaped to the image dimension $H \times W$, resulting in input channels with the same value at each position for each direction. These are concatenated together with the segmentation predictions $\mathbf{m}'$ and the corresponding image $\mathbf{x}$. We refer to our architecture as SA-SonoNet, a shape- and spatially-aware SonoNet.

Clinical ultrasound images have embedded text and calipers. While these do not impact segmentation performance, they can introduce confounding factors when training a classification model [18]. To reduce their effect on generalization, confounders are removed by applying thresholding in hue space to identify text and using the telea inpainting method [22] to eliminate calipers from the input image $\mathbf{x}$ before feeding it into our model. Our framework is summarized in Fig 1.

Existing work often faces evaluation limitations, such as small datasets and lack of method comparison, calibration evaluation, and model transparency [28]. In this section, we address these issues by conducting a comprehensive comparison of the proposed method with current approaches on a large dataset.

Dataset. Our main dataset comprises 7862 transvaginal ultrasound images extracted from a national fetal ultrasound screening database (ANONYMIZED). The original images have different resolutions with the same physical pixel distance in each direction which varies in the range $[0.037, 0.276]$ with a mean of 0.116 mm and standard deviation of 0.027 mm. The samples cover a range of gestational age (GA) at scan time ranging from week 19 to 32, with equal representation of preterm/term births per GA week. To ensure robust evaluation, we employ a 5-fold stratified cross-validation strategy, where folds have non-overlapping patients and evenly sampled term/preterm births per GA week. Furthermore, we divide (using the same strategy) each fold into equal-size validation and test sets and swap them during assessment, resulting in 10 splits with 6290/786/786 samples for training/validation/test sets.

DTU-Net is trained and tested on an external multi-class segmentation dataset with L=14 structures which includes 908/155 cervix images for training/test and additional standard planes (1481/271 head, 892/240 abdomen, 639/129 femur).

Validation of DTU-Net and CL Estimates. As Fig. 3 illustrates, DTU-Net accurately identifies the CL. Evaluation against expert CL measurements on 155 test images shows a mean absolute error of 1.79 mm, and CL predictions are robust across scan time gestational ages (GAs). Example expert annotations, DTU-Net predictions, and CL estimations are shown in Fig. 2.

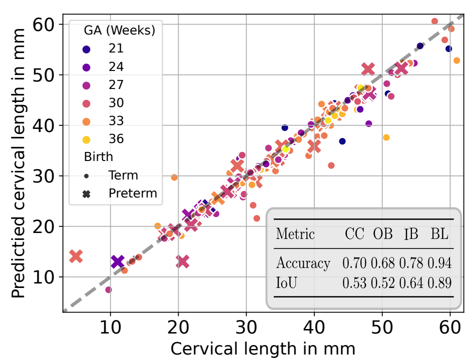


Fig. 3: CL estimates from segmentation. The table shows DTU-Net's segmentation performance in detecting the four cervical structures, CC, OB, IB, and BL (K=5, including background). Note that while accuracy and IoU are naturally low for thin curvilinear structures, the CL mean absolute error of 1.79 mm indicates that the CC segmentations are indeed appropriate for robustly estimating CL.

Baselines. We benchmark our model against 4 baseline methods, including the current clinical standard which defines sPTB when CL<25 mm. To measure CL, we employ our automated strategy, (Section 3.1). Additionally, we implement a texture-based method inspired by [3,6,11]. We extract 102 hand-crafted textural features from a binary mask covering the area near the cervical canal (see Fig 2)

Table 1: Classification results averaged across 10 balanced test splits. Reported metrics are area under the curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), and an unbiased calibration error (UCE) [19].

Method	AUC ↑	ACC ↑	SEN ↑	SPE ↑	UCE ↓
CL < 25mm	0.673 ± 0.032	0.626 ± 0.026	0.406 ± 0.043	$\mathbf{0.846 \pm 0.018}$	-
TextureNet	0.685 ± 0.025	0.642 ± 0.025	0.545 ± 0.037	0.740 ± 0.027	$\mathbf{0.015 \pm 0.016}$
MT U-Net [25]	0.700 ± 0.020	0.645 ± 0.019	0.558 ± 0.043	0.732 ± 0.023	0.079 ± 0.024
SonoNet w/PS	0.700 ± 0.032	0.645 ± 0.030	0.590 ± 0.048	0.700 ± 0.021	0.021 ± 0.024
SA-SonoNet	$\mathbf{0.750 \pm 0.034}$	$\mathbf{0.686 \pm 0.033}$	$\mathbf{0.629 \pm 0.037}$	0.743 ± 0.041	0.035 ± 0.021

and apply principal component analysis (PCA) maintaining 97% of the information, i.e, 32 PCA features, which are used to train a two-layer multi-layer perceptron (MLP) called TextureNet. Furthermore, we compare our results with a multi-task U-Net (MT U-Net) [25], trained both for classification and segmentation of the same binary mask. Finally, we include as a baseline, a pre-trained SonoNet-32 [4] injected only with pixel spacing information (SonoNet /w PS).

Implementation Details. All models were trained using binary cross-entropy loss, except MT U-Net, which followed the multitask loss defined in [25]. We used SGD optimizer with a momentum of 0.9 and batch size of 64, while we decayed the initial learning rate of 10^{-3} by a factor of 75% when the validation loss plateaued for 10 epochs. We applied an L2-regularization of 10^{-4} and saved models with the best validation loss for evaluation. TextureNet was implemented with a two-layer MLP with 128 and 64 neurons including batch normalization and drop-out layers. We used pyfeats [12] to extract early and late texture features. Our model was based on pre-trained SonoNet-32 [4], with modifications to the first layer to match the input channel size. The images were resized to 224×288, pixel intensity was normalized to $[-1, 1]$, pixel spacing was calculated

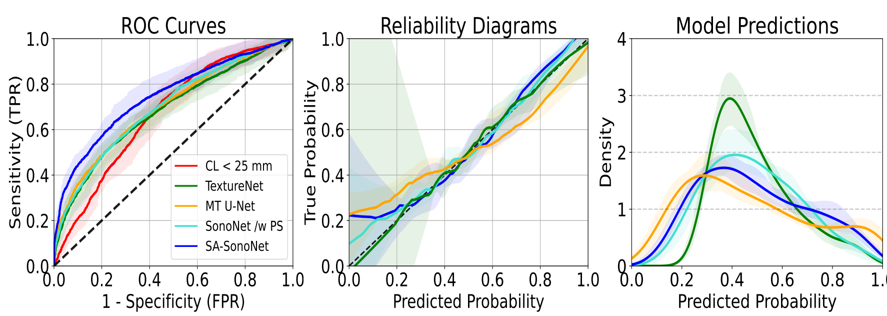


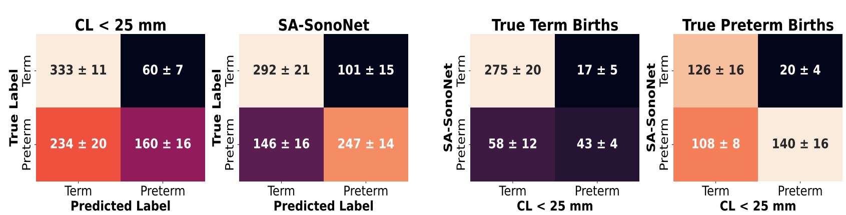
Fig. 4: **From left to right:** ROC curves, loess-based reliability diagrams, and model predictions. Uncertainty is estimated with a 95% confidence interval.

Table 2: SA-SonoNet performance when removing or modifying parts of the inputs; image (IM), cervical canal (CC), outer boundary (OB), inner boundary (IB), bladder (BL), background (BG) defined as the image with all segmented structures subtracted, and pixel spacing (PS). R stands for replacing the original PS by random sampling from pixel spacing distributions, and B stands for blacking out from the image the pixels surrounding the cervix.

IM	CC	OB	IB	BL	BG	PS	AUC	ACC	SEN	SPE
✓	✓	✓	✓	✓	✓	✓	**0.750 ± 0.034**	**0.686 ± 0.033**	0.629 ± 0.037	0.743 ± 0.041
✓						✓	0.514 ± 0.038	0.505 ± 0.023	0.771 ± 0.199	0.232 ± 0.212
✓	✓					✓	0.514 ± 0.038	0.507 ± 0.031	0.788 ± 0.190	0.220 ± 0.209
✓		✓				✓	0.549 ± 0.033	0.527 ± 0.023	0.744 ± 0.216	0.304 ± 0.249
✓			✓			✓	0.582 ± 0.048	0.539 ± 0.031	0.696 ± 0.198	0.376 ± 0.239
✓				✓		✓	0.512 ± 0.038	0.503 ± 0.025	0.780 ± 0.192	0.220 ± 0.204
✓					✓	✓	0.629 ± 0.025	0.531 ± 0.028	**0.917 ± 0.052**	0.114 ± 0.096
✓	✓				✓	✓	0.663 ± 0.015	0.576 ± 0.032	0.831 ± 0.057	0.320 ± 0.114
✓		✓			✓	✓	0.684 ± 0.021	0.615 ± 0.025	0.717 ± 0.062	0.513 ± 0.108
✓			✓		✓	✓	0.705 ± 0.041	0.611 ± 0.044	0.820 ± 0.036	0.401 ± 0.106
✓				✓	✓	✓	0.633 ± 0.027	0.537 ± 0.030	0.911 ± 0.051	0.163 ± 0.100
✓	✓	✓	✓	✓	✓		0.542 ± 0.076	0.507 ± 0.037	0.831 ± 0.173	0.185 ± 0.185
✓	✓	✓	✓	✓	✓	R	0.681 ± 0.021	0.629 ± 0.025	0.498 ± 0.096	**0.758 ± 0.060**
	✓	✓	✓	✓	✓	✓	0.705 ± 0.028	0.642 ± 0.032	0.677 ± 0.071	0.605 ± 0.104
B	✓	✓	✓	✓	✓	✓	0.733 ± 0.027	0.663 ± 0.027	0.649 ± 0.091	0.676 ± 0.103

for the resized images, and injected in mm. During training, we applied standard data augmentation such as random flipping, rotations, contrast, and brightness. Our source code for our benchmark, including preprocessing, model training, and testing, is available at github.com/ppegiosk/SA-SonoNet-sPTB.

Results. We evaluate all methods across 10 test splits in terms of area under the curve (AUC), accuracy (ACC), sensitivity (SEN), specificity (SPE), and an unbiased calibration error (UCE) [19]. Results are found in Table 1. We also provide receiver operating characteristic (ROC) curves, loess-based reliability di-



(a) Confusion Matrix (b) Confusion Matrix (c) Method Agreement (d) Method Agreement

Fig. 5: Confusion and method (dis-)agreement matrices across 10 test splits.

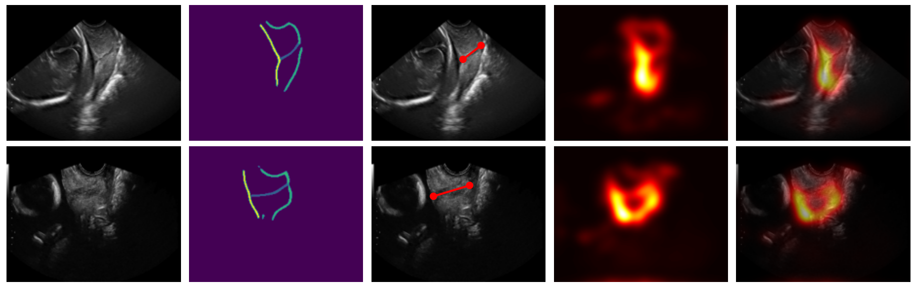


Fig. 6: Examples of high-confidence preterm birth correct predictions for a short, CL=19.8 mm, (top) and a larger, CL=30.1 mm, (bottom) cervix. **From left to right:** image without confounders, DTU-Net's segmentation predictions, automatic CL measurements, saliency maps, saliency maps on top of the input image.

agrams and distribution of model predictions in Fig. 4. Our model demonstrates competitive performance across metrics while being well-calibrated.

Feature Relevance. We assess feature relevance of different inputs of our model by evaluating model performance when removing or modifying parts of the input channels at test time. Results are found in Table 2.

Model Interpretation. Since our model is well-calibrated and performs well, confidence scores can be interpreted as risk of preterm birth. Moreover, our feature importance study brings insight into the driving features of our predictions.

For additional insight into how the model differs from the clinical standard, we provide confusion matrices for CL-based predictions in Fig. 5a and our model in Fig. 5b. Moreover, (dis)-agreement matrices between the two approaches for term and preterm births are shown in Fig. 5c and Fig. 5d, respectively.

Finally, we leverage the explainability of SonoNet [4], which is a fully CNN architecture replacing standard fully-connected layers with 1×1 convolutions. This modification maintains the correspondence between class score maps and input images, enabling the extraction of class-specific saliency maps. Examples of high-confidence predictions are shown in Fig. 6.

4 Discussion and Conclusion

The experimental results highlight the effectiveness of our method for predicting sPTB by incorporating cervical shape and spatial information into our model. Our automated CL estimation approach achieves similar predictive performance for sPTB prediction as previous studies [5,6] that relied on manual CL measurements. Our model achieves the highest specificity beyond the current clinical standard while maintaining the highest sensitivity among all methods. The calibration evaluation demonstrates the reliability of our model's confidence scores, enabling their interpretation as risks of preterm birth in a potential clinical setting. Our feature importance study shows the importance of segmentation and pixel spacing inputs in driving model predictions, enhancing transparency. Furthermore, Fig. 5d shows that our model identifies pregnancies at risk in cases of CL$>$25 mm where the current clinical standard falls short. Leveraging its inherent architecture, we generate saliency maps that enhance model interpretability.

Limitations. Our model assumes a well-trained segmentation model. Obtaining expert annotations for segmentation models can pose a practical constraint. However, we show the feasibility of using a DTU-Net [16] trained for general multi-class segmentation tasks and more standard planes than the cervix.

Summary. In this paper, we propose SA-SonoNet, a shape- and spatially-aware network for predicting sPTB. Differing from existing methods, SA-SonoNet leverages segmentation predictions and pixel spacing as additional information. We validated our model on a large class-balanced clinical dataset consisting of 7862 images, where the proposed method surpasses baselines considerably.

Acknowledgements. This work was supported by the Pioneer Centre for AI (DNRF grant nr P1), the DIREC project EXPLAIN-ME (9142-00001B), the Novo Nordisk Foundation through the Center for Basic Machine Learning Research in Life Science (NNF20OC0062606) and SONAI, an AI signature project from the Danish Agency for Digital Government.

References

1. Ahmed, W.S., Madny, E., Habash, Y., Ibrahim, Z., Morsy, A., Said, M.: Ultrasonographic wall thickness measurement of the upper and lower uterine segments in the prediction of the progress of preterm labour. Clinical and Experimental Obstetrics & Gynecology **42**(3), 331–335 (2015)
2. Akazawa, M., Hashimoto, K.: Prediction of preterm birth using artificial intelligence: a systematic review. J Obstetrics and Gynaecology **42**(6), 1662–1668 (2022)
3. Baños, N., Perez-Moreno, A., Julià, C., Murillo-Bravo, C., Coronado, D., Gratacos, E., Deprest, J., Palacio, M.: Quantitative analysis of cervical texture by ultrasound in mid-pregnancy and association with spontaneous preterm birth. Ultrasound in Obstetrics & Gynecology **51**(5), 637–643 (2018)
4. Baumgartner, C.F., Kamnitsas, K., Matthew, J., Fletcher, T.P., Smith, S., Koch, L.M., Kainz, B., Rueckert, D.: SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. IEEE TMI **36**(11) (2017)
5. Burgos-Artizzu, X.P., Baños, N., Coronado-Gutiérrez, D., Ponce, J., Valenzuela-Alcaraz, B., Moreno-Espinosa, A.L., Grau, L., Perez-Moreno, Á., Gratacós, E., Palacio, M.: Mid-trimester prediction of spontaneous preterm birth with automated cervical quantitative ultrasound texture analysis and cervical length: a prospective study. Scientific Reports **11**(1), 1–7 (2021)

6. Bustamante, D., Yan, Y., Basij, M., Gelareh, A., Hernandez-Andrade, E., Shams, S., Mehrmohammadi, M.: Cervix ultrasound texture analysis to differentiate between term and preterm birth pregnancy: A machine learning approach. In: IEEE IUS. pp. 1–4. IEEE IUS (2022)
7. Coutinho, C., Sotiriadis, A., Odibo, A., Khalil, A., D'Antonio, F., Feltovich, H., Salomon, L., Sheehan, P., Napolitano, R., Berghella, V., et al.: Isuog practice guidelines: role of ultrasound in the prediction of spontaneous preterm birth. Ultrasound in obstetrics & gynecology: the official journal of the International Society of Ultrasound in Obstetrics and Gynecology **60**(3), 435–456 (2022)
8. Dagle, A.B., Liu, Y., Crosby, D., Feltovich, H., House, M., Yan, Q., Myers, K.M., Jambawalikar, S.: Automated segmentation of cervical anatomy to interrogate preterm birth. In: Perinatal, Preterm and Paediatric Image Analysis: 7th International Workshop, PIPPI 2022, Held in Conjunction with MICCAI 2022, Singapore, September 18, 2022, Proceedings. pp. 48–59. Springer (2022)
9. Dziadosz, M., Bennett, T.A., Dolin, C., Honart, A.W., Pham, A., Lee, S.S., Pivo, S., Roman, A.S.: Uterocervical angle: a novel ultrasound screening tool to predict spontaneous preterm birth. Am J Obstetrics and Gynecology **215**(3) (2016)
10. Farràs Llobet, A., Regincós Martí, L., Higueras, T., Calero Fernández, I.Z., Gascón Portalés, A., Goya Canino, M.M., Carreras Moratonas, E.: The uterocervical angle and its relationship with preterm birth. J Maternal-Fetal & Neonatal Medicine **31**(14), 1881–1884 (2018)
11. Fiset, S., Martel, A., Glanc, P., Barrett, J., Melamed, N.: Prediction of spontaneous preterm birth among twin gestations using machine learning and texture analysis of cervical ultrasound images. U Toronto Medical Journal **96**(1) (2019)
12. Giakoumoglou, N.: Pyfeats (2021). https://doi.org/10.5281/zenodo.6783286
13. Hemming, V.G., Overall Jr, J.C., Britt, M.R.: Nosocomial infections in a newborn intensive-care unit: results of forty-one months of surveillance. New England Journal of Medicine **294**(24), 1310–1316 (1976)
14. Hermann, K., Chen, T., Kornblith, S.: The origins and prevalence of texture bias in convolutional neural networks. NeurIPS **33**, 19000–19015 (2020)
15. Lin, M., Feragen, A., Bashir, Z., Tolsgaard, M.G., Christensen, A.N.: I saw, i conceived, i concluded: Progressive concepts as bottlenecks. arXiv:2211.10630 (2022)
16. Lin, M., Zepf, K., Christensen, A.N., Bashir, Z., Svendsen, M.B.S., Tolsgaard, M., Feragen, A.: Dtu-net: Learning topological similarity for curvilinear structure segmentation. In: International Conference on Information Processing in Medical Imaging. pp. 654–666. Springer (2023)
17. Luechathananon, S., Songthamwat, M., Chaiyarach, S.: Uterocervical angle and cervical length as a tool to predict preterm birth in threatened preterm labor. Int J Women's Health pp. 153–159 (2021)
18. Mikolaj, K., Lin, M., Bashir, Z., Svendsen, M.B.S., Tolsgaard, M., Nymark, A., Feragen, A.: Removing confounding information from fetal ultrasound images. arXiv:2303.13918 (2023)
19. Petersen, E., Ganz, M., Holm, S., Feragen, A.: On (assessing) the fairness of risk score models. In: FAccT. pp. 817–829 (2023)
20. Pizzella, S., El Helou, N., Chubiz, J., Wang, L.V., Tuuli, M.G., England, S.K., Stout, M.J.: Evolving cervical imaging technologies to predict preterm birth. In: Seminars in Immunopathology. vol. 42, pp. 385–396. Springer (2020)
21. Sepúlveda-Martínez, A., Diaz, F., Muñoz, H., Valdés, E., Parra-Cordero, M.: Second-trimester anterior cervical angle in a low-risk population as a marker for spontaneous preterm delivery. Fetal diagnosis and therapy **41**(3), 220–225 (2017)

22. Telea, A.: An image inpainting technique based on the fast marching method. Journal of graphics tools **9**(1), 23–34 (2004)
23. Ven, V.D., et al.: The capacity of mid-pregnancy cervical length to predict preterm birth in low-risk women: a national cohort study. Acta obstetricia et gynecologica Scandinavica **94**(11), 1223–1234 (2015)
24. Vogel, J.P., Chawanpaiboon, S., Moller, A.B., Watananirun, K., Bonet, M., Lumbiganon, P.: The global epidemiology of preterm birth. Best Practice & Research Clinical Obstetrics & Gynaecology **52**, 3–12 (2018)
25. Włodarczyk, T., Płotka, S., Rokita, P., Sochacki-Wójcicka, N., Wójcicki, J., Lipa, M., Trzciński, T.: Spontaneous preterm birth prediction using convolutional neural networks. In: Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis: First International Workshop, ASMUS 2020, and 5th International Workshop, PIPPI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4-8, 2020, Proceedings 1. pp. 274–283. Springer (2020)
26. Włodarczyk, T., Płotka, S., Szczepański, T., Rokita, P., Sochacki-Wojcicka, N., Wojcicki, J., Lipa, M., Trzciński, T.: Machine learning methods for preterm birth prediction: a review. Electronics **10**(5), 586 (2021)
27. Włodarczyk, T., Płotka, S., Trzciński, T., Rokita, P., Sochacki-Wójcicka, N., Lipa, M., Wójcicki, J.: Estimation of preterm birth markers with u-net segmentation network. In: Smart Ultrasound Imaging and Perinatal, Preterm and Paediatric Image Analysis: First International Workshop, SUSI 2019, and 4th International Workshop, PIPPI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 4. pp. 95–103. Springer (2019)
28. Yang, Q., Fan, X., Cao, X., Hao, W., Lu, J., Wei, J., Tian, J., Yin, M., Ge, L.: Reporting and risk of bias of prediction models based on machine learning methods in preterm birth: A systematic review. Acta Obstetricia et Gynecologica Scandinavica **102**(1), 7–14 (2023)
29. Zuo, J., McFarlin, B.L., Simpson, D.G., O'Brien, W.D., Han, A.: Automated region of interest placement on cervical ultrasound images for assessing preterm birth risk. The Journal of the Acoustical Society of America **153**(3), A352–A352 (2023)