Received: 27 March 2023 | Revised: 5 June 2023 | Accepted: 17 July 2023

DOI: 10.1002/pd.6411

REVIEW

PRENATAL DIAGNOSIS WILEY

Artificial intelligence in obstetric ultrasound: A scoping review

Rebecca Horgan | Lea Nehme | Alfred Abuhamad

Division of Maternal Fetal Medicine, Department of Obstetrics & Gynecology, Eastern Virginia Medical School, Norfolk, Virginia, USA

Correspondence
Rebecca Horgan, Division of Maternal Fetal Medicine, Department of Obstetrics & Gynecology, Eastern Virginia Medical School, Norfolk, VA, USA.
Email: horganr@evms.edu

Abstract

The objective is to summarize the current use of artificial intelligence (AI) in obstetric ultrasound. PubMed, Cochrane Library, and ClinicalTrials.gov databases were searched using the following keywords "neural networks", OR "artificial intelligence", OR "machine learning", OR "deep learning", AND "obstetrics", OR "obstetrical", OR "fetus", OR "foetus", OR "fetal", OR "foetal", OR "pregnancy", or "pregnant", AND "ultrasound" from inception through May 2022. The search was limited to the English language. Studies were eligible for inclusion if they described the use of AI in obstetric ultrasound. Obstetric ultrasound was defined as the process of obtaining ultrasound images of a fetus, amniotic fluid, or placenta. AI was defined as the use of neural networks, machine learning, or deep learning methods. The authors' search identified a total of 127 papers that fulfilled our inclusion criteria. The current uses of AI in obstetric ultrasound include first trimester pregnancy ultrasound, assessment of placenta, fetal biometry, fetal echocardiography, fetal neurosonography, assessment of fetal anatomy, and other uses including assessment of fetal lung maturity and screening for risk of adverse pregnancy outcomes. AI holds the potential to improve the ultrasound efficiency, pregnancy outcomes in low resource settings, detection of congenital malformations and prediction of adverse pregnancy outcomes.

Key points

What is already known about this topic?

- The development of artificial intelligence (AI) in obstetric ultrasound is currently in its infancy, as fetal ultrasound poses a number of unique challenges including the mobility of the fetus, the developing fetal anatomy, the requirement to obtain specific planes for diagnosis, which can be both difficult to obtain and limited by fetal positioning and maternal body habitus.

What does this study add?

- This is the first scoping literature review of AI in obstetric ultrasound. We have provided a comprehensive overview of the current capabilities, challenges and potential future uses of AI in obstetric ultrasound.

- AI holds the potential to improve the ultrasound efficiency, pregnancy outcomes in low resource settings, detection of congenital malformations and prediction of adverse pregnancy outcomes.

1 | INTRODUCTION

Antenatal ultrasound examination is the primary mode of imaging in pregnancy for assessment of the fetus. Accurate diagnosis of major fetal malformations, growth disorders and placental abnormalities allows for timely and appropriate pregnancy management. Obstetric ultrasound has several challenges, including its operator dependency, and its steep learning curve. Furthermore, access to adequately trained personnel and equipment is limited in low resource settings.

Artificial intelligence (AI) is the ability of computer programs to perform processes associated with human intelligence, such as reasoning, learning, adaptation, sensory understanding and interaction.[1] Machine learning is a set of powerful computational tools that train models on descriptive patterns obtained from human inference rules.[2] A major limitation of machine learning is that it relies heavily on statistical insights and thus can be resource intense, requiring labeling of large volumes of images to train a model. Deep learning is a branch of machine learning that utilizes convolutional neural networks (CNNs). CNNs use principles from linear algebra to provide a scalable approach to image classification and object recognition and have the ability to attain a high level of performance with limited training samples.[2]

The use of AI in the field of radiology has considerably developed in recent years, particularly in the diagnosis of liver, thyroid and breast diseases.[3–7] The development of AI in obstetric ultrasound is currently in its infancy, as fetal ultrasound poses a number of unique challenges; the mobility of the fetus, the developing fetal anatomy, the requirement to obtain specific planes for diagnosis which can be both difficult to obtain and limited by fetal positioning and maternal body habitus. Fetal factors such as speckle noise, occlusion of boundaries and other artifacts can also affect intelligent detection and measurement.[8–11] We chose to conduct a scoping review of AI in obstetric ultrasound due to the heterogeneity of available studies, to allow us to identify research conduct in the field, and to identify current knowledge gaps. Our primary objective was to summarize the current use of AI in obstetric ultrasound.

2 | METHODS

This scoping review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement specific to scoping reviews.[12] A literature search was conducted using Pubmed, Clinicaltrials.gov and the Cochrane library from inception through May 2022. Studies eligible for inclusion were identified using the following search strategy: "neural networks" OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "transformer model" AND "obstetrics" OR "obstetrical" OR "fetus" OR "foetus" OR "fetal" OR "foetal" OR "pregnancy" or "pregnant" AND "ultrasound." The search was limited to the English language. In addition, the reference section of each included article was reviewed to assess additional articles eligible for inclusion.

Two authors (R.H., L.N.) independently screened all abstracts to assess eligibility for inclusion. Studies deemed potentially relevant were then full-text reviewed by both authors (R.H., L.N.). Studies were eligible for inclusion if they described the use of AI in obstetric ultrasound. For the purpose of this scoping review, obstetric ultrasound was defined as the process of obtaining ultrasound images of a fetus, amniotic fluid, or placenta. AI was defined as the use of neural networks, machine learning, or deep learning methods. All study types were eligible for inclusion, including observational studies, cohort studies, randomized control trials, quantitative, qualitative and mixed-methods studies. Exclusion criteria included expert opinions or review articles, use of semi-automated systems requiring physician or sonographer input, and studies which described machine-learning methods to analyze numeric data previously obtained using ultrasound.

Data extraction was performed using a standardized data abstraction tool designed for this study. For all studies eligible for inclusion, the following data were extracted: number of included study participants, inclusion criteria, primary outcome, the type of AI used, the description of AI method and summary of results. Data were tabulated using Microsoft Excel and summarized using a narrative review and descriptive statistics.

3 | RESULTS

A total of 127 papers fulfilled our inclusion criteria (Figure 1: PRISMA flowchart). Eleven studies focused on AI in the first trimester (Table 1). Three studies focused on the automated detection and measurement of nuchal translucency.[13–15] Three studies focused on the detection and measurement of the gestational sac in the first trimester.[16–18] Three studies focused on the detection of the midsagittal plane of the fetus, with one study focusing on 2D ultrasound[19] and two studies on 3D ultrasound.[20,21] One study focused on three-dimensional ultrasound in the first trimester to measure fetal biometry and to detect the presence of fetal limbs[22] and the final study focused on segmentation and measurement of the cerebral cortex from 2D images.[23]

Eight studies focused on AI for assessment of the placenta (Table 2). Three studies focused on deep learning methods for automatic segmentation of the placenta, with two of these studies

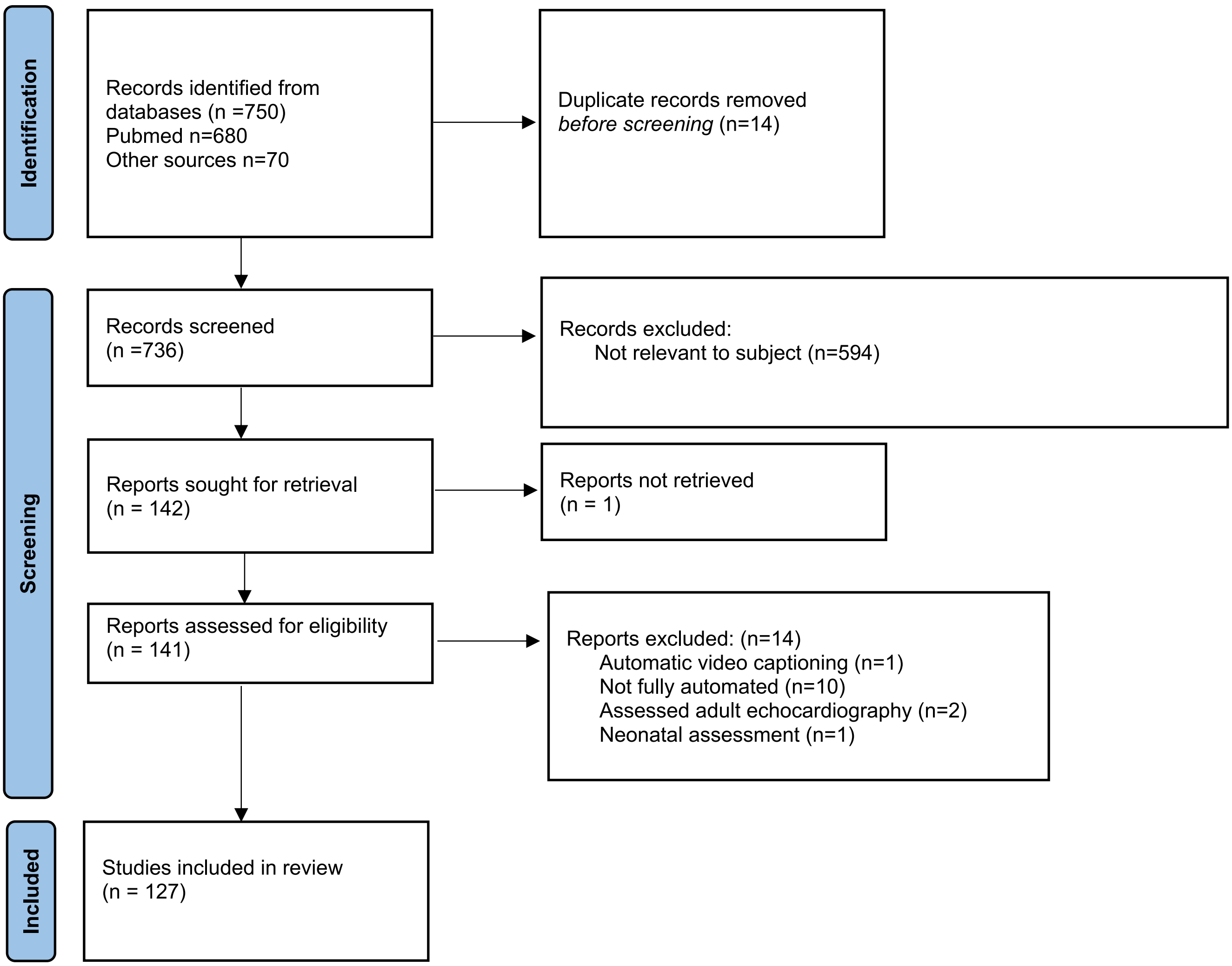

FIGURE 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. [Colour figure can be viewed at wileyonlinelibrary.com]

focusing on 3D ultrasound data.[24–26] Looney et al. evaluated the clinical utility of placental volume by assessing first trimester placental volumes for the prediction of small for gestational age (SGA) neonates at birth.[24] The authors found that the receiver-operating characteristics curves for the log placental volume (MoMs) calculated by the fully automated convoluted neural network (OxNNet) and the real world technique to predict SGA were almost identical at 0.65 (95% CI 0.61–0.69) for OxNNet and 0.65 (95% CI 0.61–0.70) for the ground-truth.[24] One study each focused on categorization of the placenta as normal, low-lying or placenta previa,[27] automatic detection of placental location from ultrasound sweeps obtained from non-specialist sonographers,[28] automated staging of placental maturity[29] and automated detection of placental lacunae.[30] The final study by Gupta et al. compared placental quantitative image texture throughout pregnancy in patients with hypertensive disorders to normotensive pregnancies with normal outcome, with placentas categorized from placental images using deep learning methods.[31] The authors found that sensitivity and specificity for abnormal placental image texture were 70.6% and 76.6% in the first trimester, 60.4% and 73.3% in the second trimester, and 83.5% and 83.5% in the third trimester for the development of hypertensive disorders.

Forty-seven studies focused on the use of AI for the assessment of fetal biometry (Table 3). Ten studies assessed automated measurement of the fetal abdominal circumference (AC), femur length (FL), head circumference and biparietal diameter.[8,32–40] Arroyo et al. focused on deep learning to assess fetal presentation and placental location, in addition to assessment of fetal biometry for estimation of gestational age in the third trimester.[33] Twenty studies focused on the automation of fetal head measurements only with automated detection of the correct scanning plane and automated measurements of various head and intracranial structures.[9–11,41–57] Studies by both Pluym and Grandjean used 3D head volumes for obtaining automated measurements. One study assessed automated analysis of fetal brain morphology on standard cranial ultrasound sections to estimate the gestational age in second and third trimester fetuses, compared with standard fetal biometry using a CNN with supervised learning based on previously labeled images.[58] Five studies assessed deep learning for automated measurement of the FL only[59–63] and two studies assessed automated measurement of AC only.[64,65] There

TABLE 1 First trimester studies.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Gofer 2021	56 fetuses	12–14 weeks, no major CNS anomalies on 1st trimester ultrasound	Two image segmentation methods processed high-resolution fetal brain images obtained during the NT scan: "Statistical region Merging" and "Trainable Weka Segmentation" measurement of the fetal cerebral cortex in original and processed images served to evaluate the performance of the algorithms	To evaluate the feasibility of ML tools for segmenting and classifying first-trimester fetal brain ultrasound images	A mean absolute percentage error of fetal cerebral cortex measurement of 1.71% ± 0.59 SD was observed and performance was similar in fetuses with normal versus abnormal NT measurements
Tsai 2020	218 patients	11–13 weeks, healthy patients, singleton pregnancies	The authors developed a system using deep learning in 2 stages The 1st stage is to find a seed point for mid sagittal point followed by the use of generative adversarial network for mid sagittal plane detection in 3D ultrasound scans. The authors compared an automated and semi-automated system to manual measurements by an expert	To propose a framework that would allow the automated precise detection of middle sagittal plane on 3D ultrasound	The four metrics exhibited no significant differences in five-fold cross-validation when comparing automated and manual assessment. In the automatic system results, 98.6% (n = 215) had Euclidean distances <0.05. The authors found that the automatic system was two times faster than a semi-automated approach
Ryou 2019	65 patients	First trimester (11–14 weeks), healthy pregnancies	A deep learning and image processing method was developed for segmentation of the fetus, to detect plane orientation, to localize and estimate fetal biometry and to identify fetal limbs	To propose a new automated system for first trimester 3D ultrasound	This automated method came close to human expertise in 3D ultrasound assessment
Sciortino 2017	12 patients, 382 random frames	11–13 weeks	Wavelet and multi-resolution analyses were used for detection of the NT region through identification of the jawbone and then measurement of the NT	To propose a method for automatic detection and measurement of the NT	True positive rate: 99.95% for nuchal area detection and 64% for measurements
Nie 2016	346 3D ultrasound images from 204 patients	11–13 + 6 weeks	A deep belief network was built to detect the fetal head from 3D US data using an enhanced circle detection method for detection of the size and position of the fetal head. A model was then constructed with six parameters for the detection of the sagittal plane	To propose an automatic technique for detecting the sagittal plane on 3D ultrasound	The results showed that the plane had a small distance error of <4 mm 88.6% of the time and an angle error of <20° in 71% of the time with overall accuracy of 91.62%
Sciortino 2015	3000 frames	11–13 weeks	A wavelet analysis and neural network was developed to detect the jawbone and radial analysis was developed to identify the choroid plexus to enable detection of the mid-sagittal plane	To propose a new method for automated identification of the mid-sagittal sections in first trimester ultrasound	The results were as follows (%): True positive: 87.26 True negative: 94.98 Accuracy: 91.12

(Continues)

TABLE 1 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Park 2013	196 DICOM scans	1st trimester	An algorithm was created to locate the fetal head through learning detectors and to estimate the NT region. An NT cut is refined and measured using Dijkstra's shortest path applied on an edge-enhanced image	To propose a fully automated method for measuring NT in the midsagittal plane	This method is both effective and efficient in detecting and measuring NT. Only the results of the 5 worst NT detection cases and 5 best NT detection cases were shown. The first 5 had an average error of 0.24 mm and the latter 5 had an average error of 0.29 mm
Deng 2012	690 ultrasound images	NR	Three classifiers were trained through Gaussian pyramids to detect the NT region. A spatial model was proposed to define spatial constraints and dynamic programming for the appropriate detection of the NT	To propose a hierarchical model for the automated detection of NT, fetal head and body	The accuracy of NT with 50% of training data was 55.9%
Zhang 2012	92 pregnant women	Singleton pregnancies	Adaboost classifiers were trained for efficient detection of the GS and a snake model was then used to segment and measure GS	To propose a method to decrease interobserver variability related to localization of the early GS and performance of GS biometric measurements	The differences between system performance and radiologist performance with respect to GS selection and length and depth measurements were $7.5 \pm 5.0\%$, $5.5 \pm 5.2\%$, and $6.5 \pm 4.6\%$ respectively
Zhang 2011	31 videos for testing and 61 images for training	NR	A model was trained using a database and Adaboost algorithm to locate and measure the GS	To propose a 3 stage method to locate the GS	Standardized plane of GS error is 1 with average measurement error of 0.059 cm for length diameters and 0.083 cm for depth
Borenstein 2009	65 pregnancies	11–13 + 6 weeks healthy fetuses	This study measured the GS volume excluding the fetus and the placenta using VOCAL software and then using SonoAVC software to compare results. SonoAVC is sonography based automated volume count uses 3D ultrasound initially developed to measure follicular volume. VOCAL is virtual organ computer-aided analysis and is a used to perform manual volume calculations	To assess the SonoAVC system accuracy in measuring GS volume and to compare with the VOCAL system	In 95% of cases, SonoAVC was able to measure the volume of GS and the success rate increased with GA. VOCAL and SonoAVC systems had comparable results

Abbreviations: 3D, three dimensional; CNS, central nervous system; GS, gestational sac; ML, machine learning; NT, nuchal translucency.

were four studies assessing correct detection of the fetal abdominal standard plane (FASP) using AI[66–69] and five studies assessed various combinations of two or more fetal biometry measurements.[70–74]

Ten studies focused on the use of AI in fetal cardiac imaging (Table 4). Sakai et al. developed a deep learning-based explainable representation to visualize the detection of substructures of the heart in a 2D screening video and calculated an abnormality score by measuring the deviation from normal. This allowed the ultrasound examiner to use the graph chart diagram and abnormality score to perform fetal cardiac ultrasound screening and the authors demonstrated significant improvements in the detection of congenital heart disease.[75] Abuhamad et al. evaluated a system that automatically

TABLE 2 Placental studies.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Gupta 2021	429 patients	First trimester 11–14 weeks	Images of the placenta were taken serially in the first, second, and third trimester, processed and classified using validated deep learning tools. The authors used five pretrained models "wide_resnet50_2," "wide_resnet101_2," "resnext50_32×4d," "resnext101_32×8d," "googlenet" and used image data augmentation techniques to artificially expand the size of a training data-set by creating modified versions of images in the dataset	To compare placental quantitative image texture in patients with hypertensive disorders to those with normal outcome	The image texture disparity between cases and controls was highly significant ($p < 0.001$). The model "resnext 101_32×8d" had Cohen kappa score of 0.413 (moderate) and the accuracy score of 0.710 (good). In the first, second and third trimester sensitivity and specificity for abnormal placental image texture were 70.6% and 76.6%, 60.4% and 73.3%, and 83.5% and 83.5% respectively
Schilpzand 2021	280 pregnant women	18–40 weeks	Deep learning using U-Net for segmentation of the placenta, and then categorization of the placenta as normal, low-lying or previa	To propose a deep learning method that would allow the detection of low-lying or placenta previa from 2D ultrasound images	Test dice coefficient of 0.84 (IQR + 0.23) Sensitivity 81% Specificity 82%
Torrents-Barrena 2021	60 images	17–37 weeks, singleton and mono- di twin pregnancies	Thirteen state-of-the-art 3D networks were examined, to determine the optimal architectural components for the segmentation of placenta. All networks used deep learning segmentation methods to learn features automatically from the raw data in an end-to-end fashion	To evaluate several state-of-the-art deep learning-based segmentation approaches to automatically segment the placenta from 3D US data	With regards to placenta detection in 3D US data, all networks perform similarly with the exception of HolisticNet, DenseASPP, and DeepMedic (Jaccard from 0.41 to 0.57). The best Dice values range from 0.70 ± 0.06–0.76 ± 0.12
Saavedra 2020	10 pregnant women	Age: 22–43 years, 3rd trimester	Images were collected using volume sweep imaging by a non specialist and U-Net deep learning was used to segment the placenta. The output masks combined with the knowledge of the acquisition protocol assessed the spatial location of the placenta (using a heat map)	To propose an automatic method for the detection of the location of the placenta	The method showed a sensitivity of 75% and a specificity of 92% for placental location
Hu 2019	1364 fetal ultrasound images from 247 pts	8–34 weeks, singleton and multiple gestation	The authors developed a framework using a convolutional neural network that includes a layer weighted by automated acoustic shadow detection	To develop a new method for placental segmentation using convolutional neural network	Mean dice coefficients for automated segmentation on the full dataset with and without the acoustic shadow detection layer were 0.92 ± 0.04 and

(Continues)

TABLE 2 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
					0.91 ± 0.03 when compared to manual segmentation
Looney 2018	2393 first trimester 3D-US volumes	11 + 0–13 + 6 weeks	A fully convolutional neural network (OxNNet) was trained to automatically segment the placenta using 3D US volumes. The training data set was quality controlled by three operators to produce the "ground-truth" data set	To fully automate the segmentation of the placenta from 3D ultrasound volumes and to assess the clinical utility of placental volume by assessing predictions of small-for-gestational-age babies at term	The ROC curves for the log placental volume (MoMs) calculated by the fully automated CNN (OxNNet) and the real world technique to predict SGA were almost identical at 0.65 (95% CI 0.61–0.69) for OxNNet and 0.65 (95% CI 0.61–0.70) for the ground-truth
Qi 2018	34 patients including 23 with invasive placentas and 11 with noninvasive placentas	NR	A layered aggregation structure based on deeply supervised IDA for automatic placental lacunae localization was proposed by the authors	To design a model to automatically detect placental lacunae in 2D placental ultrasound images	The model yielded the highest mean average precision of 35.7%, surpassing all other baseline models evaluated (32.6%, 32.2%, 29.7%)
Lei 2015	443 images	18–40 weeks	A supervised learning method using a support vector machine was used to extract features, cluster features using Gausian mixture model and encode the clusters by Fisher vector for staging accuracy enhancement	To propose a method to detect and automatically stage placenta maturity from B mode ultrasound images	The model demonstrated an AUC of 96.77%, sensitivity: 98.04%, specificity: 93.75%, 98.04%, and 93.75% for the placental maturity staging

Abbreviations: 2D, two dimensional; 3D, three dimensional; AUC, area under the curve; CNN, convoluted neural network; IDA, iterative deep aggregation; MoM, multiples of the median; NR, not reported; ROC, receiver-operating characteristics; SGA, small for gestational age; US, ultrasound.

retrieves the right and left ventricular outflow tracts from a three-dimensional (3D) volume of the fetal chest and found they were correctly identified 91.7% and 94.4% of the time, respectively.[76] Arnaout et al. used trained neural networks to detect the various cardiac views and distinguish between normal and abnormal images on fetal echocardiogram.[77] The authors found an AUC of 0.99, a sensitivity of 95%, and specificity of 96%. Herling et al. evaluated the correlation between automated measurement of fetal atrioventricular plane displacement using myocardial velocity traces obtained by color tissue Doppler imaging (cTDI) versus those obtained by anatomic M mode and found a significant correlation between mitral annular plane systolic excursion ($r = 0.64$; $P < 0.001$), septal annuar plane systolic excursion ($r = 0.72$; $P < 0.001$) and tricuspid annular plane systolic excursion ($r = 0.84$; $P < 0.001$) measurements obtained by M-mode and those obtained by cTDI.[78] Dozen et al. developed a novel method for image segmentation of ultrasound videos based on deep learning on the four-chamber view of the fetal heart to accurately assess the ventricular septum and reported a mean Intersection over Union of 0.5543.[79] Three studies proposed automated systems for detection and evaluation of the fetal four-chamber view whereas Dong et al. proposed a deep learning framework for quality control of the four-chamber view, demonstrating a mean average precision of 94.52%.[80–83] A study by Xi et al. proposed a system for automatic segmentation of the fetal heart and lungs from ultrasound images.[84]

Twenty studies focused on the use of AI for fetal neurosonography beyond the first trimester (Table 5). Three studies evaluated deep learning algorithms to localize planes within the fetal brain from 3D ultrasound volumes,[85–87] whereas four additional studies proposed models to segment or measure various intracranial structures from 3D US volumes of the fetal head.[88–91] Two studies focused on deep learning systems for the detection of fetal intracranial planes from 2D ultrasound.[92,93] Nambuerete et al. developed a system to predict gestational age and neurodevelopmental maturation of the fetus from 3D images.[94] Five studies developed models for automatic detection or measurement of various fetal intracranial structures from 2D images in the second trimester including the lateral ventricles, cerebellum, cavum septum pellucidum, corpus

TABLE 3 Fetal biometry.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Arroyo 2022	58	Third trimester, maternal age ≥18	A standardized volume sweep image was performed in the absence of a trained sonographer and a deep learning algorithm (U-Net) was used for diagnostic assessment	To assess the accuracy of U-Net to determine fetal position, placental location and fetal biometric measurements compared to an experienced sonographer	U-Net showed 100% agreement for fetal presentation (Cohen's κ 1 [$p < 0.0001$]) and 76.7% agreement for placental location (Cohen's κ 0.59 [$p < 0.0001$]). It also achieved relative error of 5.6% for BPD and 7.9% for HC. Biometry measurements corresponded to estimated gestational age within 2 weeks of those assigned by standard of care examination with 89% accuracy
Chenarlogh 2022	HC-18 dataset 1334 2D images MFP dataset: 999 images	NR	The authors used a deep learning method, U-Net, which includes three major blocks: The feature encoder path, the bottleneck layer, and the feature decoder path. They then used their dataset to test HC and AC segmentation by U-Net	To propose a new fast and accurate U-Net model for medical imaging segmentation	Dice and Jaccard coefficients were 97.62% and 95.43% for fetal head segmentation and 95.07%, and 91.99% for fetal abdominal segmentation. Dice and Jaccard coefficients of 97.45% and 95.00% using the public HC18-Grand challenge dataset
Plotka 2022	50	19–38 weeks gestation	A novel multi-task convolutional neural network based spatio-temporal fetal US extraction and standard plane detection algorithm (FUVAI) was created by the authors and evaluated on 50 freehand fetal US video scans. The FUVAI obtained measurements were then compared to measurements obtained by five experienced sonographers	To investigates the use of deep convolutional neural networks to automatically perform measurements of fetal HC, AC, BPD, FL using fetal ultrasound videos to obtain EFW and estimate GA	FUVAI had similar performance to the sonographer obtained measurements and operates within the range of human-level error. The estimated GA and fetal weight has a mean absolute error of 0.05 ± 0.01 weeks and 25 ± 5 g respectively when comparing FUVAI obtained measurements to the experienced sonographers
Wang 2022	551 pregnant women	HC-18 database	GAC Net is a CNN that is based on an encoder decoder system and can perform end-to-end training. The attenuation model (SUO) was then used	To propose a convolutional neural network for the accurate detection and measurement of HC	GAC-Net had the following results: HD: 1.22 ± 0.71 mm Absolute difference: 1.75 ± 1.71 mm DSC: 98.21 ± 1.16% of head circumference

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			to identify the correct structural boundaries		which were superior to other deep learning models including U-Net, V-Net and Mask-RCNN
Zhang 2022	551 pregnant women	HC-18 database	*Segmentation based model*: The use of convolutional neural networks for medical image segmentation, processing of segmentation results to exclude all inaccurate results and finally HC computation through ellipse fitting *Segmentation free model*: The use of regression convolutional neural networks which can learn the features of the fetal head and estimate the head circumference directly	To compare segmentation based and segmentation free methods for measurement of fetal HC	The results show that segmentation based approaches provide more accurate results, however the learning that is possible with regression CNN models may result in better estimations of HC in the future
Bano 2021	349 US images from 42 pregnancies	NR	The authors developed a framework that used semantic segmentation models to segment key anatomical features as well as region fitting and finally scale recovery for the biometry estimation	To propose an automated framework for measurements required for calculating EFW	The authors noted the system was accurate for biometry estimation with the error between clinically measured and predicted biometry less than the permissible error during clinical measurements. Comparison of the predicted versus clinically measured fetal biometry showed that the errors in HC (0.67 mm), AC (3.77 mm) and FL (2.10 mm) were minimal and better than the ±15% error that is typically accepted in fetal US assessment
Burgos-Artizzu 2021	1992	Second or third trimester singleton pregnancy with no congenital malformation or aneuploidy	A novel method for GA estimation (quantusGA) from the TTA plane was developed using deep learning techniques based on CNN. The	To evaluate the performance of quantusGA on automated analysis of fetal brain morphology on standard cranial	95% confidence interval of the error in gestational age estimation was 14.2 days for the artificial intelligence method alone and

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			method is based on supervised learning of previously labeled images	ultrasound sections to estimate the gestational age in second and third trimester fetuses compared with standard fetal biometry	11.0 days when used in combination with fetal biometric parameters, compared with 12.9 days using standard biometrics alone
Luo 2021	1005	Low risk pregnancy, between 16 and 41 weeks, singleton, accurately dated by 1st trimester US, no structural anomalies	The authors used SF to acquire sections and measure biparietal diameter, head circumference, abdominal circumference and femur length. SF consists of two applications: SFA and SFM. SFA is a technique in which only one finger touch is used during real-time scanning and automatically distinguishes acquired standard sections in the cine loop that contain the specific standard section and then automatically measures related growth parameters	To evaluate the efficacy of SF in standard obtaining biometric measurements compared to traditional ultrasound	In 998 of 1005 cases (99.30%), SF successfully acquired the sections and made all measurements. The agreement between the techniques was high for all measurements. The time to obtain sections and measure biometric parameters or solely measure biometric parameters was significantly shorter with SF than with traditional ultrasound. The authors concluded that SF helped in the acquisition of reliable standard sections and biometric measurements and saved time
Moccia 2021	551 women	NR	A convolutional neural network, Mask-R2CNN, was developed using the data from the HC18 Grand challenge with 999 images used for training and 335 images for testing on delineating HC by regressing distance fields	To present a deep learning approach that can accurately delineate HC in fetal ultrasound images	Mask-R2CNN was able to address the challenges of HC delineation in US images, with an absolute difference of 1.95 mm, without any manual intervention
Oghli 2021	999 images for training and 335 for testing	14–26 weeks GA	Attention MFP-Unet learns to detect and extract the anatomical structures that are of interest through a convolutional neural network	To propose a convolutional neural network for segmentation of fetal biometric parameters and measuring BPD HC AC and FL	Superior performance of attention MFP-Unet compared to other approaches used for automatic measurement of fetal biometric parameters was noted with the following results: DSC: 0.98 HD: 1.14 mm Good contours: 100% Conformity: 0.95

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Pluym 2021	143 women	18–22 + 6 weeks GA	A sonographer and a physician obtained manual measurements using 2D ulltrasound for BPD, HC, transcerebellar diameter, cisterna magna, and posterior horn of the lateral ventricle. Then 3D imaging sweep was obtained by the sonographer and automated measurements taken by SonoCNS fetal brain. SonoCNS was developed by GE healthcare it utilizes a 3D sweep of the fetal brain to align three views into a single view in which it recognizes and measures the structures	To compare the accuracy of an automated 3D ultrasound technique with regards to fetal intracranial measurement to the traditional manual technique	The ICC reflected a reliability of 0.8–0.88 for the automated approach compared to the manual measurements for BPD and HC and a poor-moderate reliability (0.23–0.5) for transcerebellar diameter, cisterna magna and posterior horn of the lateral ventricle
Prieto 2021	23,309 images from ZAPPS study and 124,645 images from UNC used to train the generative networks. 7233 images from FAMLI study used to evaluate AI system	NR	A large database of obstetric US images acquired, stored and annotated by expert sonographers was used to train algorithms to classify, segment, and measure several fetal structures: BPD, HC, CRL, AC and FL. Raw images were then used for model training by removing caliper and text annotation to fully automate image classification, segmentation, and structure measurement to estimate the GA	To assess automatic measurement of fetal structures using a low-cost obstetric US to assist in establishing GA without the need for skilled sonographer	There was an average accuracy of 93% in classification tasks, a mean intersection over Union accuracy of 0.91 during segmentation tasks, and a mean measurement error of 1.89 cm, leading to a 1.4 days mean average error in the predicted GA compared to expert sonographer GA estimate using the Hadlock equation
Zeng 2021	551 pregnant women	HC-18 database	Data in the HC-18 database was used to train DAG V-Net deep learning models. These models were able to segment fetal head scans from 2D US images. Ellipse fitting was then used for automated measurement of HC	To propose a deep learning technique for segmentation of fetal ultrasound scans and measurement of HC	DSC: 97.93% DF (HC difference): 0.09 ± 2.45 mm AD (absolute difference): 1.77 ± 1.69 mm HD: 1.29 ± 0.79 mm This method ranked 5th in the HC18 challenge

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Zhu 2021	435 images	NR	The authors compared traditional machine learning method to determine the edges of the femur based on a forest regression method and automatic measurements of FL based on a deep learning method (SegNet)	To evaFplotate and clinically validate and the proposed algorithm presented for measuring fetal femur length	Compared to manual performance of FL, the approach based on the forest regression model (traditional machine learning method) was 1.23 ± 4.66 mm and SegNet was 0.46 ± 2.82 mm
Fiorentino 2020	335 images	NR	A region-proposal CNN for head localization and centering was created followed by a regression CNN for accurately delineating the HC. The first CNN was trained to exploit transfer learning, while the regression CNN was based on distance fields	To lower intra-and inter-operator variability in HC measurements	A mean absolute difference of 1.90 (±1.76) mm and a DSC of 97.75 (±1.32) % were achieved
Li 2020	551 patients 1334 images	Pregnant women receiving routine US screening with clinically healthy fetuses	An automatic measurement system that is based on a neural network which provides information for fetal head segmentation, accurate BPD and OFD prediction	To present a novel end to end deep learning network to measure HC, BPD and OFD automatically from 2D images	SAPNet had an overall average mean intersection over Union accuracy of 96.46 ± 1.77
Miyagi 2020	Japanese Society of Ultrasonics in Medicine dataset	18–41 weeks gestation	Neural networks were trained by deep learning to estimate fetal weight based on the gestational age, BPD, AC and FL	To develop an AI method to estimate fetal weight based on BPD, AC and FL	The authors concluded that the AI's good accuracy for extreme fetal weights is likely to be very useful and that AI with the neural network seems to have potential for estimating fetal weights
Zhang 2020	HC-18 dataset	All trimesters of pregnancy	The authors used a CNN with the ability to learn on its own to detect the contour of the head using 4 regression models	To propose a new method where the CNN can directly measure HC without manually labeling segmented images	The deeper model Reg-ResNet50 had better performance with mean squared error loss function than the rest of the models
Al Bander 2019	999 2D images for training and 335 images for testing from 551 pregnant women	NR	The authors proposed a framework to detect the fetal head using object localization and segmentation based on a deep learning model. They used a fully convolutional neural network developed	To propose a deep learning based method to segment the fetal head from US images	The authors found the method efficient for fetal head biometry measurement with a Dice coefficient of 97.73 ± 1.32

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			for segmentation to improve the accuracy of the image and ellipse fitting at the contour of the head for measurement		
Rajinikanth 2019	HC-18 dataset: 999 images for training and 335 for testing	NR	The authors used a Jaya algorithm and Otsu threshold for pre-processing of 2D US images and Chan Vese and Level set segmentation for post-processing. The final step was then evaluation and validation of the system	To use a hybrid-scheme to determine HC on 2D US images	The hybrid procedure offers enhanced picture similarity during HC measurement (>88.5%) compared to 2D ultrasound images and the ground truth which are annotations on the existing dataset
Rong 2019	1334 images for fetal HC	NR	CNNs are used to derive an external force which is integrated into active contour models. This method is tested on fetal head in ultrasound images. Active contour models fit an ellipse at the fetal head boundaries	To propose a new algorithm that trains a convolutional neural network to derive an external force. This external force is integrated in the active contour models for curve evolution and ellipse fitting of the fetal head circumference	This method is effective in detecting fetal head circumference and it is competitive to other methods used for this task ADF (absolute difference): 2.45 ± 2.55 DSC: 95.49 ± 4.11 HD: 2.44 ± 1.96
Salim 2019	100 patients	Singleton gestation, 20–40 weeks, normal BMI, age 18–35 pregnancy dating by CRL	An algorithm was used to detect HC, AC and FL, and measurement performed by automatically inserting calipers. The caliper placement was assessed subjectively and classified as acceptable, minor changes required, major changes required. The authors compared the automatic measurements with manual measurements	To evaluate whether an automated method can detect and measure HC, AC and FL on 2D US images	This automated tool identified the correct biometric measurements in 99% of the images. The results were accurate compared to manual measurements
Sobhaninia 2019	999 images	All trimesters included	The authors used a multi-task deep convolutional neural network for automated segmentation of the HC and an Ellipse tuner for measuring HC	To propose a new method for the automated measurement of fetal HC	The authors demonstrated a DSC score 96.84 ± 2.89, comparable with the radiologist's annotations

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Van Den Heuvel 2019	183 pregnant women	NR	The obstetric sweep protocol was used to identify head circumference measurements using two fully convolutional neural networks VGG net would identify the frames with fetal head and Unet would then measure the HC from the images identified by VGGnet	To develop a system that can directly measure fetal head circumference by utilizing data from obstetric sweep protocol	The HC measurements were used to estimate gestational age through the Hadlock equation Most of the results were between P 2.5 and P 97.5 intervals of the Hadlock curve [P 2.5–P 97.5 interval is a 95% percent prediction interval where P is percentile]
Grandjean 2018	30	16–30 weeks gestation, maternal age >18, singleton	Smartplanes: AI software that enables the automatic identification of correct scanning plane within the head volume and automatic positioning of calipers to measure BPD and HC	To evaluate the feasibility and reproducibility of Smartplanes to automatically identify the TTA plane from 3D ultrasound volumes to measure BPD and HC compared to manually obtained measurements by 2 experienced sonographers	Interclass correlation coefficients were >0.9 for comparison between 3D measurements obtained by Smartplanes software and the sonographer measurements
Kim 2018	172 images	NR	The authors developed a deep learning process that was hierarchical and divided into 3 steps: 1. Measurements of HC and BPD (ellipse fitting) 2. Plane acceptance check 3. Refinement of the measurements	To develop a deep learning system to estimate HC and BPD with high accuracy and reliability	The authors demonstrated a success rate of 92.31% for HC and BPD accuracy of 87.14% for the acceptance of the plane
Kim 2018	77 pregnant women	NR	The authors used a CNN, U-Net for initial estimation of AC, measurement of AC and plane checking. These processes take into account clinicians' decisions, anatomical structures and characteristics of the ultrasound image	To propose an automated method for fetal biometry estimation	U-Net demonstrated: Dice similarity metric of 92.55 $\pm$ 0.83 for AC measurement Accuracy of 87.10% for acceptance check of the FASP
Sinclair 2018	2724 2D images	18–22 weeks	The authors assessed convolutional networks that are trained to segment fetal head ultrasounds on 2D images and used ellipse fitting for the	To propose an automated system to estimate measurements of HC and BPD	HC: Model expert error: 1.99 mm Inter-observer error 2.16 mm BPD: Model expert error: 0.61 mm

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			segmented contour to measure HC and BPD		Inter-observer error 0.59 mm Dice coefficient: Model expert error: 0.980 Inter-observer error 0.980 The authors concluded model performance resembles human expert performance
Li 2017	145 images of HC for testing and 524 for training.	18–33 weeks gestation	The framework integrated knowledge about GA and US images into a random forest classifier and then used phase symmetry to detect the center of the fetal skull. Finally, the ellipse fitting method (ElliFit) was used to measure HC	To propose a learning-based framework that used knowledge and ElliFit to automatically measure fetal HC	The authors found that the framework had an average measurement error of 1.7 mm for fetal HC
Jang 2017	56 training cases 32 test cases	NR	The author's developed a convolutional neural network that uses doctor's experience, anatomical structures and the characteristics that define a certain ultrasound image and then used Hough transformation to measure the AC	To propose a model for automated measurement of fetal AC from 2D ultrasound data	CNN accuracy compared to expert: 0.809 CNN accuracy compared to expert 2: 0.771 Accuracy between 2 experts 0.905
Wu 2017	492 US videos for training and 219 for testing	16–40 weeks	FUIQA uses two deep convolutional neural networks (L-CNN and C-CNN) L-CNN finds the region of interest and C-CNN assesses the quality of the image by assessing stomach bubbles and umbilical vein	To propose a fetal ultrasound image quality assessment (FUIQA) for US image quality control	The authors found FUIQA results are comparable to manual assessments by experts
Wu 2017	Training set: 900 fetal head images & 688 fetal abdomen images Testing set: 236 fetal head & 505 fetal abdomen images	19–40 weeks	A cascaded CNN was utilized for feature extraction from US images and to distinguish the anatomy. The authors then used the auto-context scheme to improve the CNN	To propose a cascaded framework for automatic US image segmentation of fetal head and abdominal scans	The authors noted that large variations of GA, size, appearance and shape in anatomy were well addressed by the system on boundary delineation tasks and concluded that the CNN showed promising segmentation accuracy

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Zhang 2016	60 fetal images	20–35 weeks gestation	The authors used a deep learning method which used a filter bank designed to extract texton features related to the BPD, HC, OFD and FL. The texton cues with multiscale local brightness form a unified framework for the delineation of fetal head and femur and use square ellipse fitting for measurement of the fetal head and closed contour for the fetal femur	To propose a texton approach for segmentation of the BPD, HC, OFD, FL	Fetal head: Precision: 96.85% Maximum symmetric contour distance: 1.46 mm Average symmetric contour distance: 0.53 mm Fetal femur: Precision: 84.37% Maximum symmetric contour distance: 2.72 mm Average symmetric contour distance: 0.31 mm
Chen 2015	300 videos for training and 219 for testing	18–40 weeks	The authors developed a multi-layered CNN that was trained from Image-Net detection data to locate the FASP. The system used a classifier to generate a probability map and identify the US image as FASP or not FASP	To present a learning based approach to identify the FASP using CNN	The authors reported the following results for their CNN: Accuracy: 0.904 Precision: 0.908 Recall: 0.995 *F*1: 0.950
Hur 2015	39 pregnancies	Singleton uncomplicated pregnancies between 26 + 0 and 32 + 0 weeks	The image of the long bone was reconstructed using the 5D LB (five dimensional long bone). After the 3D volume data were displayed in a multiplanar mode The length of the bone was measured automatically.	To determine the feasibility of 5D LB functions by comparing the biometric data using 2D ultrasound, 3D volume data and 5D LB treated 3D volume	5D LB is reproducible and comparable with conventional 2D and 3D ultrasound techniques for fetal long bone measurement The interclass correlation coefficient for femur, tibia, and fibula was 0.91, 0.92, and 0.89, respectively
Perez-Gonzalez 2015	23 images	NR	The authors used a framework that incorporates texture maps, morphological operations, active contours and optimal ellipse for detection, segmentation and measurement of BPD and HC	To propose a fully automated method to segment and measure fetal head from 2D images	Precision: 94.61% Dice similarity index: 97.19% Max distance: 2.64 mm Correlation: 99.8%
Foi 2014	90 2D images	21, 28, and 33 weeks	The authors developed a framework is based on the cost function that assumes that the fetal head is elliptical.	To propose a fully automated method to segment the fetal head and measure	The segmentation accuracy was similar to the inter expert variability and better

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			A template image is then constructed and the cost function compares this template image with the observed US image. A Nelder-Mead algorithm was used for minimization	BPD and HC from 2D US images	than other automated methods The biometric measurements were as accurate as the manual measurements Dice %: 97.73 ± 0.89 Maximum symmetric contour distance: 2.26 ± 1.47 mm Average symmetric contour distance: 0.91 ± 0.47 mm
Rueda 2014	90 images	Fetuses at 21, 28 and 33 weeks	The authors evaluated various methods to automatically segment fetal anatomy to measure standard planes for the fetal head and femur at different gestational ages (21, 28 and 33 weeks) with varying image quality to reflect data encountered in real life environments. Five teams completed in the challenge and experts manually delineated the objects of interest to define the ground truth	To evaluate and compare segmentation methods for measurement of biometry	The authors found that head sub-challenge resulted in better results than femur sub-challenge and it was comparable to manual measurements
Wang 2014	90 fetal ultrasounds	Fetuses at 21, 28 and 33 weeks	The authors used a fully automatic deep learning system to segment the femur from US images and measure fetal length	To propose an automatic method for fetal femur segmentation and measurement	The authors concluded that the method works well at extracting the femur and measurement of the femur length with the system taking an average of 2.28 s Maximum symmetric contour distance: 6.02 ± 7.29 Average symmetric contour distance: 1.04 ± 1.29 Root mean square symmetric contour distance measurement: 1.77 ± 2.41
Chen 2012	300 videos for training and 219 for testing	18–40 weeks gestation	The authors developed a deep convolutional neural network and a transfer learning	To propose a learning based approach to detect FASP	This T-CNN method was shown to be better than R-CNN (regional based) and RVD

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			strategy constructed by US videos to locate the FASP		(radial component model and vessel probability map) with: Accuracy: 0.896 Precision 0.714 Recall: 0.710 *F*1: 0.712
Ponomarev 2012	90 fetal femur and 90 fetal head images	NR	The framework uses the relative brightness of the femur and head to locate the region of interest and then uses segmentation and multilevel thresholding on these planes. A support vector machine classifier was used for the selection from the segmented images of a valid scan of the femur and an edge-based scoring function was used for refinement of the skull ellipse	To propose a fully automated method for the detection of the standard fetal planes for biometric measures from 2D images	93.4% of the images of the femur were correctly segmented and measure and 96.6% of the fetal head images were correctly segmented and measured
Sun 2012	90 images	NR	The authors developed an algorithm containing the circular shortest path extraction, robust ellipse fitting and skull outer edge finding to automate measurements	To propose an algorithm to automatically measure BPD, OFD and HC	The authors found the circular shortest path algorithm both automatic and efficient. The typical running time for the algorithm on a 756 $\times$ 546 image is about 1–2 s
Rahmatullah 2011	2384 images	14 weeks-term gestation, healthy women, singleton pregnancies with no fetal abnormalities	Adaboost learning algorithm was trained to extract images containing the two landmarks (stomach bubble and umbilical vein)	To propose an automated method to detect two landmarks: The stomach bubble and umbilical vein	Detection of stomach bubble was more accurate than umbilical vein detection with an execution time <6 s. The umbilical vein detection was poor at early gestational ages
Mukherjee 2010	90 images	Third trimester fetuses	The algorithm has a normalized score based on the size, shape and presentation of the femur in clinically acceptable scans A polynomial curve fitting technique is used for the delineation of the end points of the femur to allow it to be measured	To propose an automated two step framework for detecting and measuring FL	The predictions from automated measurements were found to be within $\pm$2 SD of GA estimates from both manual measurements in 89/90 cases and were within $\pm$3 SD in all 90 cases

(Continues)

TABLE 3 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Shrimali 2009	7 subjects	Non-anomalous singleton, live born fetuses with no family history of dwarfism; no maternal use of alcohol or cigarettes or maternal diabetes	Morphological operators processed the images to remove background, then refined them followed by measurement of the FL. The whole process was done in a time efficient manner	To propose a new time efficient morphology-based algorithm to detect refine and measure femur length	Time for detection, refinement and measurement was 4 s. The authors found the results were comparable to manually measured FL
Carneiro 2007	1426 head, 1168 femur, and 1293 abdomen images for training Test set: 177 head, 183 abdomen, and 171 femur images	NR	The authors developed a deep learning method derived from a constrained probabilistic boosting tree and a large database of expert annotated images for detection and segmentation of the HD, BPD, AC and FL	To propose a new technique for fast automatic fetal biometry measurements	This method was efficient and closely accurate to experts in obstetric measurements with average error of 0.0265 with respect to ground truth measurements
Carneiro 2007	1426 head images 1168 femur 1293 abdomen 547 humerus 325 fetal body	NR	The authors evaluated a deep learning technique for automatic segmentation of fetal biometric measurements in addition to fetal humeral length and CRL, using constrained probabilistic boosting tree and expert annotated images	To propose an efficient, robust and accurate technique for segmentation of biometric measurement as well as HL and CRL	The authors noted that the system had similar accuracy to measurements by experts with results as follows: CO of BPD: 0.71 mm (σ 0.61), 1.19% (σ = 0.85) IO of BPD: 0.83 mm (σ 0.66), 1.33% (σ 0.82) *r* for BPD: 0.999 CO of HC: 5.22 mm (σ 5.27), 2.07% (σ = 1.67) IO of HC: 8.46 mm (σ 3.28), 3.54% (σ 0.99) *r* for HC: 0.996 CO of AC: 12.6 mm (σ 9.48), 6.35% (σ 5.26) IO of AC: 11.62 mm (σ 10.6), 5.65% (σ 6.53) *r* for AC: 0.974 [CO = mean computer to observer distance, IO = mean inter observer difference and *r* is the correlation coefficient]

Abbreviations: 2D, two dimensional; 3D, three dimensional; AC, abdominal circumference; AI, artificial intelligence; BPD, biparietal diameter; CNN, convolutional neural network; CRL, crown rump length; DSC, dice similarity coefficient; EFW, estimated fetal weight; FASP, fetal abdominal standard plane; FL, femur length; GA, gestational age; HC, head circumference; HD, hausdorff difference; HL, humeral length; ICC, intraclass correlation coefficients; NR, not reported; OFD, occipitofrontal diameter; SD, standard deviation; SF, smart fetus; SFA, smart fetus acquisition; SFM, smart fetus measurement; T-CNN, tube convolutional neural network; UNC, University of North Carolina; US, ultrasound.

TABLE 4 Fetal echocardiography.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Primary objective	Results
Sakai 2022	160 cases 344 videos	18–34 weeks	This interpretable model is an auto-encoder that includes two novel techniques, cascade graph encoder and view-proxy loss, and generates a "graph chart diagram" as an explainable representation. The graph chart diagram visualizes the detection of substructures of the heart and vessels in the screening video on a two-dimensional trajectory and, thereafter, calculates the abnormality score by measuring the deviation from the normal. The examiner uses the graph chart diagram and abnormality score to perform fetal cardiac ultrasound screening	The authors proposed a novel deep learning-based explainable representation "graph chart diagram" to support fetal cardiac ultrasound screening, which has low detection rates of congenital heart diseases	Graph chart diagrams improved detection of abnormalities as shown by the mean AUC of the ROC curve Residents: 0.616→0.748 Fellows: 0.829→0.890 Experts: 0.966→0.975
Arnaout 2021	107,823 images from 1326 echocardiograms and 4108 fetal surveys	18–24 weeks GA	Neural networks were trained to detect the various cardiac views and distinguish between normal and abnormal CHD. Segmentation was also used to measure fetal cardiothoracic space	To propose a segmentation modality capable of distinguishing normal from abnormal echocardiograms	This model had: AUC: 0.99 Sensitivity: 95% Specificity: 96% Negative predictive value: 100% in differentiating between normal and abnormal fetal echocardiograms
Herling 2021	Group 1 = 201 Group 2 = 107 Group 3 = 35	Group 1: Uncomplicated singleton gestation at 18–42 weeks used to develop reference ranges, Group 2: Uncomplicated singleton gestation at >41 weeks gestation, Group 3: EFW <2.5 centile or <10th centile with UA PI >97.5th centile	Cineloops of the heart were obtained using cTDI and an automated algorithm developed in-house was used to obtain mitral, tricuspid and septal annular plane systolic excursion. Gestational-age specific reference ranges were constructed and normalized for cardiac size	To evaluate the correlation between automated measurement of fetal atrioventricular plane displacement using myocardial velocity traces obtained by cTDI versus those obtained by anatomic M mode	There was a significant correlation between MAPSE ($r = 0.64$; $P < 0.001$), SAPSE ($r = 0.72$; $P < 0.001$) and TAPSE ($r = 0.84$; $P < 0.001$) measurements obtained by M-mode and those obtained by cTDI
Qiao 2021	1000 images	NR	The authors proposed a PSFFGAN, which synthesizes high-quality fetal four chamber views using four chamber sketch images. In addition,	To propose a PSFFGAN that synthesizes high quality four chamber views using four chamber sketch images	The experimental results show that the fetal four views synthesized by the proposed PSFFGAN have objective evaluation values as follows:

(Continues)

TABLE 4 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Primary objective	Results
			they proposed a novel TGALF, which optimizes PSFFGAN to fully extract the cardiac anatomical structure information provided by four chamber sketch images to synthesize the corresponding fetal four chamber views with speckle noises, artifacts, and other ultrasonic characteristics		Structural similarity index of 0.4627, Multiscale structural similarity of 0.6224, and freshet inception distance of 83.92, respectively
Xi 2021	312 images	37 weeks GA	The authors developed a model to achieve semantic segmentation of the fetal heart and lungs using a multiscale model with skip connection framework and attention mechanisms integrated. The multi-scale feature extraction modules are incorporated with additive attention gate units for irrelevant feature elimination, through a U-Net framework with skip connections for information compensation	To propose an automated semantic segmentation model of fetal hearts and lungs from ultrasound images	Dice coefficients of fetal heart segmentation obtained by the authors proposed method vary in the range of 0.878–0.904 and Dice coefficients for segmentation of fetal lungs ranged from 0.784 to 0.872
Dozen 2020	211 pregnant women	Normal fetal cardiac ultrasound	The authors developed a novel method for segmenting the fetal ventricular septum called CSC, which employs the time-series information of videos and specific section information to calibrate the output of a deep learning system, U-net	To develop a novel method for image segmentation of ultrasound videos based on deep learning on the four-chamber view of the fetal heart to accurately assess the ventricular septum	The mIOU was 0.5543 for the CSC which was superior to other deep learning methods (U-Net and DeepLab v3+)
Dong 2020	2032 positive and 5000 negative images.	14–28 weeks	The authors developed a framework composed of three networks; a basic CNN, for classifying 4 chamber views from raw data, a deeper CNN to determine the gain and zoom of the images and an aggregated	To propose a deep learning framework for quality control of the four chamber view	The authors demonstrated a mAP of 93.52%. The adaptability and generalization had a mAP of 81.2%

TABLE 4 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Primary objective	Results
			residual visual block net to detect the anatomical structures on a plane. The quantitative score of all 3 networks achieves automatic quality control		
Qiao 2017	1250 images	Healthy pregnant women	The authors developed an artificial intelligent learning system incorporating a multistage residual hybrid attention module to capture discriminative features of the fetal cardiac chambers and to integrate residual identity mapping to alleviate information loss allowing the system to accurately locate the four-chamber view	To propose an intelligent feature learning detection system to automatically obtain the four chamber view in the fetus	The authors noted the following results: Precision: 0.919 Recall: 0.971 *F*1 score: 0.944 mAP 0.953 Frame per second: 43
Sundaresan 2017	12 patients 91 ultrasound videos	20–35 weeks gestation Normal fetal heart	The authors developed a semantic segmentation model using fetal cardiac ultrasound videos with four pixelwise labels: FC (four chamber), LVOT, 3V (3 vessel) and NH (non-heart, background pixels) using a 16 layer deep convolutional neural network	To propose an automated method for identifying the fetal heart and its standard viewing plane using a fully convolutional neural network	The author's model demonstrated a classification error rate of 23.48% on real-world clinical ultrasound data
Abuhamad 2008	72 fetuses	18–23 weeks normal cardiac anatomy	3D volumes of the fetal chest were acquired at the level of the four-chamber view. Tomographic ultrasound imaging was added to the display of each diagnostic plane. The left ventricular outflow plane, the right ventricular outflow plane and the abdominal circumference plane were retrieved by the software from the 3D volumes and the data were analyzed to	To evaluate prospectively the performance of software that automatically retrieves, from a three-dimensional volume of the fetal chest, three diagnostic cardiac planes in the second trimester of pregnancy	The software identified the correctly the target planes as follows: LVOT: 94.4% RVOT: 91.7% AC: 97.2%

(Continues)

TABLE 4 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Primary objective	Results
			determine whether cardiac planes 1–3 were displayed correctly in each volume		

Abbreviations: AC, abdominal circumference; AUC, area under the curve; CHD, congenital heart defects; CNN, convolutional neural network; CSC, cropping-segmentation-calibration; cTDI, color tissue Doppler imaging; EFW, estimated fetal weight; GA, gestational age; LVOT, left ventricular outflow tract; mAP, mean average precision; MAPSE, mitral annular plane systolic excursion; mIOU, mean Intersection Over Union; NR, not reported; PSFFGAN, pseudo-siamese feature fusion generative adversarial network; ROC, receiver operator curve; RVOT, right ventricular outflow tract; SAPSE, septum annular plane systolic excursion; TAPSE, tricuspid annular plane systolic excursion; TGALF, triplet generative adversarial loss function.

callosum and choroid plexus.[95–99] Lin et al. developed and validated an AI system to automatically detect nine specific intracranial abnormalities for use during real time imaging[100] whereas Xie et al. developed a system to identify five specific common brain abnormalities including hydrocephalus, ventriculomegaly, Blake's pouch cyst, Dandy Walker malformation and cerebellar vermis hypoplasia.[101] Sahli et al. assessed automated measurement of fetal head biometry and classification as normal, microcephaly or dolichocephaly.[102] The remaining studies evaluated segmentation of the fetal cranium[103] and detection and localization of abnormal lesions in the axial plane to classify the image as normal or abnormal.[104]

Six studies focused on anatomical evaluation of the fetus (Table 6). Two studies evaluated AI systems to automatically detect 13 standard anatomic planes.[105,106] In the study by Matthews et al., feedback during live scanning was provided via a traffic light system to inform the sonographer that the required anatomic views including fetal biometry had been obtained. Sharma et al. assessed the use of AI to comprehensively analyze and quantify operator clinical workflow in a spatio-temporal context during fetal morphology ultrasound to evaluate inter and intraobserver variability.[107] One study evaluated the use of an AI system incorporating an attenuation gate to increase precision of anatomic plane detection,[108] one study evaluated an AI system to classify 14 different fetal structures in 2-D fetal ultrasound images[109] and one study evaluated a CNN for the automated detection of three standard anatomical planes including the abdomen, axial plane of the face and four chamber cardiac view.[110]

Twenty-five studies were classified as other uses of AI in obstetric use of ultrasound (Table 7). Four studies evaluated AI assessment of fetal lung ultrasound for assessment of fetal lung maturity, prediction of neonatal respiratory morbidity and assessment of correlation of fetal lung texture with gestational age.[111–114] Three studies assessed the ability of deep learning models to undergo self-supervised learning by correcting the order of a reshuffled fetal video clip,[115] by context restoration of unlabeled 2D fetal images[116] and the use of random forests classifiers for classification of unlabeled fetal ultrasound images.[117] Two studies focused on automated detection of fetal facial standard planes.[118,119] Two studies focused on amniotic fluid with one evaluating automatic measurement of amniotic fluid index[120] and the other assessing segmentation of amniotic fluid and fetal tissue.[121] The remaining studies included a study to assess machine learning to determine occiput anterior versus occipitoposterior position in the second stage of labor,[122] machine learning assessment of fetal-lung texture in pregnancies affected by gestational diabetes or preeclampsia compared to normal pregnancies,[123] use of an AI classifier to recognize fetal facial expressions on 4D ultrasound,[124] detection of the FASP,[125] automatic detection of the fetal face on 3D ultrasound,[126] assessment of fetal presentation and confirmation of fetal cardiac activity,[127] segmentation of the fetal thoracic wall,[128] classification of the umbilical cord into normocoiling, hypocoiling and hypercoiling,[129] automated grading of hydronephrosis on ultrasound,[130] segmentation of the fetal kidneys,[131] segmentation of the AC and FL,[132] assessment of an image reconstruction framework applied to the whole fetus[133] and automated detection of fetal standard planes.[134,135] A summary of results is reported in Table 8.

4 | DISCUSSION

Our scoping review synthesizes the current uses of AI in obstetric ultrasound. We have demonstrated that AI has the potential to not only automate time consuming ultrasound tasks through features such as automated detection and measurement of fetal biometry from ultrasound clips but also to improve the detection of congenital anomalies. Several studies have focused on congenital heart defects (CHD) which have current prenatal detection rates ranging from 14% to 87% with significant geographical variation.[136] Prenatal diagnosis of CHD significantly improves neonatal morbidity and mortality.[137,138] The future of AI in obstetric ultrasound may be its use in conjunction with human experts as demonstrated in the study by Sakai et al., which used deep learning methods to visualize the detection of substructures of the heart in a 2D screening video and then calculated an abnormality score by measuring the deviation from normal.[75] Physicians then utilized this information when assessing the fetus and the authors demonstrated improved detection rates of CHD amongst all levels of providers including residents, fellows and experts. Arnaout et al. developed trained neural

TABLE 5 Fetal neurosonography.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Di Vece 2022	6 fetuses	21–39 weeks GA	The authors proposed a regression CNN using image features to estimate the six-dimensional pose of arbitrarily oriented US planes relative to the fetal brain center. The network was trained on images acquired from phantom 3D US volumes and fine-tuned using real ultrasound data. Training data was generated by slicing US volumes into imaging planes in Unity at random coordinates and more densely around the standard transventricular plane	To build an automated ultrasound plane localization system of fetal head images for 3D visualization, training, and guidance	With phantom data, the median errors were 0.90 mm/1.17° and 0.44 mm/1.21° for random planes and planes close to the transventricular one, respectively. With real data, using a different fetus with the same gestational age, these errors were 11.84 mm/25.17°. The authors concluded that good accuracy was achieved on phantom experiments but errors remained high on real data
Hesse 2022	278 images	18–26 weeks GA, no congenital malformations	A CNN was developed composed of multi-label 3D U-Net with batch normalization for the automated segmentation of the CP, LPVH, CSPV, and CB from 3D ultrasound	To accurately segment the CP, LPVH, CSPV, and CB in 3D US image volumes during the second trimester using a minimal number of voxel-wise annotations	The study demonstrated the feasibility of subcortical segmentation in 3D US using deep learning, and shows that volumetric measures obtained from these models can be used to obtain an improved understanding of subcortical growth during gestation
Lin 2022	16,297 pregnancies (43,890 images) 166 pregnancies (169 videos)	Normal fetuses and fetuses with CNS malformation between 18 and 40 gestational weeks	The authors developed and validated an AI system, the PAICS, to detect nine specific intracranial-malformation patterns in standard sonographic reference planes of the fetal central nervous system for use during live scanning	To develop an AI system that can detect congenital central nervous system malformations	The system was able to identify intracranial image patterns with an AUC of 0.981 (95% CI, 0.974–0.988) in the real-time scan setting The performance of the PAICS was similar to that of expert operators but required less time (0.025 s per image for PAICS vs. 4.4 s for experts <0.001)
Chen 2020	500 test images	NR	The authors proposed a computer-aided detection framework for automatic measurement of fetal LV in 2D US images. A deep convolutional network was trained on 2400 images of LVs to perform pixel-wise segmentation. Then, the number of PPC was obtained via morphological operations	To assess the reliability and efficacy of the proposed framework for automatic fetal LV measurement by comparing automated measurements to those obtained by three experienced sonographers	The system had a mean absolute error of 1.8 ± 3.4 mm for LV measurement compared to sonographers measurements

(Continues)

TABLE 5 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			guided by prior knowledge, converted to a physical length and used to determine the diameter of the LV by employing the minimum enclosing rectangle method		
Mariaci 2020	Dataset A: 5000 TC and TV images Dataset B: 3736 images	Second trimester ultrasounds	The authors developed a deep learning system to evaluate a video of the fetal head obtained by a sonographer and the system then identified and measured the TCD	To propose an algorithm that would allow the estimation of gestational age through automatic detection and measurement of the TCD	The authors found that detection of the TCD frame had a high-class probability. However, there was also a high positive rate and an underestimation of TCD by automated measurement
Skelton 2020	164 cases	18 + 6–20 + 6 weeks GA anomaly ultrasound	Two observers retrospectively reviewed standard fetal head planes against predefined image quality criteria. Each fetus had 2D manually-acquired, 3D operator-selected and 3D automatically-acquired images. The proportion of adequate images from each plane and modality, and the number of inadequate images per plane was compared for each method	This assess the image quality of standard fetal head planes automatically-extracted from 3D ultrasound fetal head volumes using a customized deep learning algorithm	The authors found that the 3D DL algorithm could automatically extract standard fetal head planes from 3D-head volumes of comparable quality to operator-selected planes. However, image quality in 3D is inferior to corresponding 2D planes, likely due to limitations with 3D-technology and acquisition technique
Qu 2020	30,000 2D images 155 fetuses	16–34 weeks	The authors evaluated a system using differential CNN and feature maps that have predefined parameters to analyze patterns of pixels and thus identifying 6 fetal brain standard planes	To propose a differential CNN to differentiate 6 fetal brain standard planes from nonstandard planes	The authors reported an accuracy of 92.93%
Montero 2021	8747 images	Images from 6 different US machines, the operators had similar experience	Two GANs were trained, one for TTA and TRV views, to help deep learning ultrasound classifiers by focusing on the anatomy of the fetal brain and distinguishing TTA from transventricular axial plane images	To evaluate the generation of synthetic ultrasound fetal brain images via GANs and to apply them to improve fetal brain ultrasound plane classification	TTA: Fréchet inception distance: 13.08 Precision: 0.6616 Recall: 0.3336 Transventricular: Fréchet inception distance: 17.4856 Precision: 0.6609 Recall: 0.2850 The authors concluded that using data generated by both GANs and classical augmentation strategies resulted in

TABLE 5 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
					increasing the accuracy and area under the curve score
Xie 2020	12,780 pregnancies (10,251 normal & 2529 abnormal pregnancies)	22 + 4–26 + 3 weeks GA	Ultrasound images were divided into 2 the training and the testing scans	To examine the ability of deep learning mechanisms to correctly identify normal and abnormal ultrasound scans of the fetal brain in axial planes	Results for the deep learning model were as follows: Segmentation precision: 97.9% Recall: 90.9% DICE: 94.1% Accuracy: 96.3% Sensitivity: 96.9% Specificity: 95.9% AUC for ROC: 0.989
			The deep learning algorithm was trained to identify the fetal brain, classify it as normal or abnormal and if abnormal, to locate the lesion		The authors found localization of lesions occurred precisely in 61.6%, closely in 24.6% and were irrelevant in 13.7%
Xie 2020	92,748 pregnant women	Singleton or twin pregnancies, 18–32 weeks GA	A classifier was trained to differentiate between normal or abnormal standard brain scans (TRV or transcerebellar). The craniocerebral regions were segmented and then distributed into 4 classes. Class activation mapping was used then to localize the abnormalities (ventriculomegaly, hydrocephalus, blake pouch cyst, Dandy Walker malformation and cerebellar vermis hypoplasia)	To develop a computer aided diagnosis algorithm to identify 5 fetal common brain abnormalities	The authors achieved a DICE score of 0.942 on craniocerebral region segmentation, an *F*1 score of 0.96 for classification and a mean IOU of 0.497 for lesion localization
Alansary 2019	72 fetuses	NR	72 fetal head US scans were randomly divided into 21 and 51 images for training and testing. The chosen anatomic landmarks were manually annotated by clinical experts using three orthogonal views	To assess the use the dual DQN for fetal brain anatomic landmark location (right and left cerebellar hemisphere, CSP)	Duel DQN achieves the best accuracy detecting the right and left CER points, while DQN performs the best for finding the CSP
Sahli 2019	86 fetuses	Group 1: Uncomplicated cases Group 2: Dolichocephaly & microcephaly	The first step includes feature extraction (uses US images of BPD, OFD and HC). The second step is feature classification (machine learning method to distinguish normal from abnormal)	To propose a computerized diagnostic method based on a SVM for fetal head morphology and classifications to categorize participants into 2 groups: Normal and affected cases	Sensitivity: 0.9236 Specificity: 0.8403 Positive predictive value: 0.8160 Negative predictive value: 0.9260 Positive likelihood ratio: 6.9027 Negative likelihood ratio: 0.0943

(Continues)

TABLE 5 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
		Both groups at 22 weeks 5 days GA			The authors concluded that SVM is effective in detecting accurate fetal head diagnostics
Huang 2018	285 fetal brain images	Healthy fetus, 20–29 weeks gestation	View-based projection networks (VP-Nets), uses three view-based CNNs, to simplify 3D localizations by directly predicting 2D projections of the key structures onto three anatomical views	To assess the use of VP-Nets to detect multiple fetal brain structures simultaneously in 3D fetal neurosonography and to perform measurements of the CER and cisterna magna. These compared to manual measurements to assess accuracy	The model achieved an IOU between prediction and annotation bounding boxes of >62% on average for identification of brain structures. On average, automatic measurements correlated well with manual ones, the reported deviation is <2.0 mm
Huang 2018	5	NR	The AI model was designed to automate structure detection and segmentation of fetal brain structures based on a region descriptor that characterizes the shape and local intensity context of different neurological structures without explicit models	To automate US segmentation of fetal brain structures (corpus callosum and CP)	The results demonstrated a high region segmentation accuracy (dice coefficient: 0.81 ± 0.06% for corpus callosum, 0.76 ± 0.08% for CP) relative to human delineation
Namburete 2018	739 (599 images for training and 140 for testing)	18–34 weeks GA healthy and growth restricted fetuses of different ethnic and geographical backgrounds.	The authors developed a deep learning system using a fully CNN which can segment and locate the fetal brain and eye sockets using both 2D and 3D images	To prove that visual interpretation of fetal brain anatomy is possible through fully CNNs that aligns 3D fetal neurosonographic images on the basis of a predefined coordinate system	Co-alignment of 140 fetal ultrasound images resulted in high brain overlap and low eye localization error
Yaqub 2018	40 fetal SU of the brain	19–24 weeks gestation	The authors used a deep learning method with the use of random forests framework and the use of a classifier to detect background, CP, PVC, CSP and cerebellum (CER)	To propose an automated method to locate fetal brain structures in 3D US	3D detection accuracy: CP: 100% PVC: 80% CSP: 90% Cerebellum: 90% Comparison between the automatic detection and the manual delineation on each 2D slice from the 3D volumes showed accuracies as follows: 92.9% CP, 91.1% PVC, 91.1% CSP, 91.9% CER
Yaqub 2017	19,838 images from 10,595 fetal anatomy scans	Fetal anomaly scans	A CNN classifier i was trained and then tested to segment the fetal head region, estimate	To propose a deep learning technique to for: 1) Identification of fetal brain	The authors demonstrated a 96.9% Dice coefficient for detection of head. The

TABLE 5 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			magnification, and assess for fetal brain symmetry	2) Detecting the structures under study 3) Learning the patterns that allow the identification of that plane	system agreed with manual assessment of head magnification in 94.1%, on brain symmetry in 83.1% and on CSP visibility in 86.8% and correct assessment of head orientation in 95.3%
Nambuerte 2014	447 participants	18–34 weeks healthy	The authors developed a framework that uses a manifold surface representation of the fetal head which allows for efficient sampling of different developmental stages. The bespoke features capture neurosonographic patterns in 3D images, and using a regression forest classifier, to characterize structural brain development both spatially and temporally to capture the natural variation	To propose an automated framework for predicting GA and neuro developmental maturation based on 3D ultrasound	Estimation of GA was accurate within 6.1 days
Sofka 2014	2384 images for training	16–35 weeks	The authors developed a novel Adaboost framework for detecting structures on 3D ultrasound based on sequential estimation techniques. They used a probabilistic model as the solution for speckle noise, signal dropout, shadows and appearance variation due to difference in GA Three planes were evaluated including the cerebellar, thalamic and ventricular and the midsagittal plane	To propose an automated fetal head and brain system for measuring structures from 3D ultrasound	The authors found an average difference between manual and automatic measurements was <2 mm
Nambuerte 2013	60 2D images for testing 10 fetuses for training	25–34 weeks	The authors evaluated a deep learning framework that includes local statistics and shape information about pixel clusters in an image to evaluates the performance of the feature that segments the cranial pixels in an ultrasound image using a random forest classifier	To propose a machine learning framework for segmentation of the fetal cranium	The authors reported a 97.2% segmentation accuracy

Abbreviations: 2D, two dimensional; 3D, three dimensional; AUC, area under the curve; BPD, biparietal diameter; CB, cerebellum; CI, confidence interval; CNNs, convolutional neural networks; CP, choroid plexus; CSP, cavum septum pellucidum; CSPV, cavum septum pellucidum et vergae; DQN, deep Q networks; GA, gestational age; GANs, generative adversarial networks; HC, head circumference; IOU, intersection over union; LPVH, lateral posterior ventricle horns; LV, lateral ventricles; NR, not reported; OFD, occipitofrontal diameter; PAICS, prenatal ultrasound diagnosis artificial intelligence conduct system; PPC, pixels per centimeter; PVC, posterior ventricle cavity; SVM, support vector machine; TCD, transcerebellar diameter; TRV, transventricular; TTA, transthalamic; US, ultrasound.

TABLE 6 Anatomy ultrasound.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Matthew 2021	23	18–20 + 6 routine fetal scans, gestational age between 20 and 24	The 13 standard planes during each scan were: Transventricular brain; transcerebellar brain; abdominal circumference; femur length view; facial profile; lips and nose; right and left outflow tract; four chamber view; three vessel trachea view; kidneys; sagittal spine; coronal spine. During AI-assisted examinations, the sonographer scanned the fetus until they were satisfied that a comprehensive visual assessment had been completed. This was in combination with confirmation that the required planes had been captured (i.e. green traffic lights by the AI system). During this scan, the automated tools capture the 13 standard views and the fetal biometry	To pilot automation of anomaly scan using AI tools and assess the efficiency and quality of fetal ultrasound in comparison with traditional manual scan	Scan times were 34.7% shorter using AI assistance. Completeness of four core fetal views: AI assisted report included 93% of the required views while the manual report had 98%. Completeness of the 13 standard views: 73% for AI and 98% for manual report. The authors found automatically extracted images had better quality than the manual scans
Schlemper 2019	2694 2D ultrasound examinations	18–22 weeks GA	The authors developed a novel AG model for image analysis that automatically learns to focus on target structures of varying shapes and sizes which can be incorporated into CNNs Attention model was used for fetal ultrasound screening of: Brain, profile, lips, abdomen, kidneys, femur, spine, 4CH, 3VV, RVOT, LVOT	To propose a novel attention gate model that is easily incorporated into segmentation and classification architectures	Precision increased by 5% for kidney, fetal profile and spine, 3% for cardiac views (4CH, 3VV) which the authors noted was an improvement in precision and recall compared to sononet
Sharma 2019	25 full length scans	18–22 weeks GA	The authors evaluated the use of AI system to comprehensively analyze and quantify operator clinical workflow in a spatio-temporal context, that is, the type, duration and sequence of scanned anatomical structures and activities, in order to explore intra- and inter-operator correlation or variability	To compare several deep learning architectures and propose methods for comprehensive 2D+ spatio-temporal description in fetal anomaly US video scans	The authors noted that automated partitioning and characterization on unlabeled full-length video scans showed high correlation ($\rho = 0.95$, $p = 0.0004$) with workflow statistics of manually labeled videos, suggesting practicality of proposed methods

TABLE 6 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Sridar 2018	4074 2D images (3109 for training and 965 for testing)	NR	The authors proposed a method to automatically classify 14 different fetal structures in 2-D fetal ultrasound images by using information from both cropped regions of fetal structures and the whole image. The model used two feature extractors by fine-tuning pre-trained CNN with the whole ultrasound fetal images and the discriminant regions of the fetal structures found in the whole image 14 structures were looked at including abdomen, arm, 3-vessel view, LVOT and RVOT, cord insertion, face, femur, humerus, foot, genitals, head, heart, kidney, leg, spine, and hand	To present a method to classify 14 different fetal structures in 2D fetal ultrasound images	The system achieved a mean accuracy of 97.05%, mean precision of 76.47% and mean recall of 75.41%. The Cohen *k* of 0.72 revealed the highest agreement between the ground truth and the proposed method. The superiority of the proposed method over the other non-fusion-based methods is statistically significant ($p < 0.05$)
Baumgartner 2017	2694 2D ultrasound exams	18–22 weeks gestation	A CNN was created to automatically detect 13 fetal standard views (in accordance with the UK mid pregnancy ultrasound guidelines) in freehand 2-D ultrasound data as well as provide a localization of the fetal structures via a bounding box	To assess the accuracy of the CNN to accurately detect the standard 2D views in real time using only weak supervision	The model achieved an average *F*1-score of 0.798 in a realistic classification experiment modeling real-time detection, and obtained a 90.09% accuracy for retrospective frame retrieval. An accuracy of 77.8% was achieved on the localization task
Chen 2017	1231 videos	18–40 weeks	The authors presented a composite neural network framework that specializes in deep CNN and recurrent neural networks that can identify the different planes from fetal US videos through multitask learning. Distinct from conventional way that devise hand-crafted visual features for detection, the framework explores in- and between-plane feature learning with a novel composite framework of the convolutional and recurrent neural networks	To propose a new model that combines multi task learning, deep learning and sequence learning (RNN) model to identify the following planes: FASP FFASP FFVSP	In the testing phase, the T-RNN identified the standard planes in <1 min from a video

Abbreviations: 2D, two dimensional; 3VV, three vessel view; 4CH, four chamber; AG, attention gate; AI, artificial intelligence; CNN, convolutional neural network; FASP, fetal abdominal standard plane; FFASP, fetal face axial standard plane; FFVSP, fetal four chamber view standard plane; GA, gestational age; LVOT, left ventricular outflow tract; RVOT, right ventricular outflow tract.

TABLE 7 Other uses of obstetric ultrasound.

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Ghi 2022	1219 women	Singleton pregnancy, ≥37 weeks, non-anomalous fetus, 2nd stage of labor, cephalic	One sonographic image of the fetal head was then acquired in an axial plane using transperineal ultrasound and saved for later offline analysis. Using the transabdominal sonographic diagnosis as the gold standard, a ML algorithm based on a pattern-recognition feed-forward neural network was trained on the transperineal images to discriminate between OA and non-OA positions	To describe a ML algorithm for the automatic recognition of fetal head position using transperineal ultrasound during the second stage of labor and to describe its performance in differentiating between occiput anterior and non-OA positions.	The ML-based algorithm correctly classified the fetal occiput position in 90.4% (357/395) of the test dataset giving an *F*1-score of 88.7% and a precision-recall AUC of 85.4%
Pradipta 2022	151 images	8–32 weeks GA	This study proposed a new feature extraction method, the UCI. The model consists of five stages: Image preprocessing, feature extraction, feature selection, oversampling data using SMOTE, and classification. Machine learning method observations were then carried out comprehensively on five based classifiers: Random forest, KNN, decision tree, SVM, Näive Bayes, and Multiclassifier	To classify the umbilical cord based on ultrasound images into normocoiling, hypocoiling and hypercoiling	Random forest and multiclassifier had the highest accuracy, precision, recall and *F* measure Random forest (SMOTE 400%): 96%, 95.3%, 96.3%, and 96%. Multi ClassifierSummer21 (SMOTE 500%): 95.2%, 93.6%, 93.3%, and 93.3%
Cho 2021	255	20–36 + 6 weeks gestation, absence of oligohydramnios or PPROM, age ≥19	A hierarchical deep-learning based method was developed, which considers clinicians' anatomical knowledge based approaches. The key step is the segmentation of the AF pocket using a deep learning network, AF-net which combined three complementary concepts: Atrous convolution, multi-scale side-input layer, and side-output layer	To compare AF-net amniotic fluid measurements to those obtained by physicians	AF-net achieved a dice similarity of 0.877 ± 0.086 for AF segmentation and achieved a mean absolute error of 2.666 ± 2.986 and mean relative error of 0.018 ± 0.023 for AFI value
Du 2021	548 fetuses	28–41 weeks GA, no chromosome abnormality or congenital	Fetal-lung image acquisition was achieved using a transverse view of the	To analyze and compare, using ultrasound-based radiomics technology, fetal-lung texture in	The overall performance of the GDM and PE prediction model was superior to that of the

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
		malformation, no steroids received, singleton, no additional maternal medical complications	fetal thorax at the level of the 4-chamber view of the heart. Using ultrasound-based radiomics technology to analyze fetal-lung texture, fetal lungs were grouped according to whether they were affected by GDM or PE to see if they could be distinguished from each other and from fetal lungs of normal pregnancies	pregnancies affected by GDM or PE compared to normal pregnancies	GA prediction model, with an area under the receiver operating characteristics curve of 0.95–0.99, sensitivity of 74.5%–91.3%, specificity of 75.7%–88.4% and accuracy of 80.6%–86.4% in the independent test set
Miyagi 2021	896 images in total	Singleton fetus, 19–38 weeks GA	4D images were collected and classified into seven categories; eye blinking, mouthing, face without any expression, scowling, smiling, tongue expulsion, yawning. A deep learning AI classifier was created that consisted of convolutional neural networks with L2 regularization to obtain the probability of predicting each category of the fetal face expression. 80% of the data was used as a training data set and the remaining 20% as a validation data set	The development of an AI classifier to recognize fetal facial expressions that are considered related to brain development	The accuracy of the AI fetal facial expression classification for the entire test data set was 0.985
Weerasinghe 2021	100 scans	<18 years old singleton 20–40 weeks	The authors assessed fully convolutional networks for automated segmentation of fetal kidneys. The authors used multiparametric input fusion including 3D B-mode and power Doppler to improve accuracy	To propose a method using fully convolutional networks for automated kidney segmentation	Early input-level fusion provided the best segmentation accuracy: Average DSC of 0.81 Hausdorff distance of 8.96 mm, an improvement of 0.06 DSC and reduction of 1.43 mm Hausdorff distance compared to baseline network
Burgos-Artizzu 2020	12,400 images from 1792 patients	NR	The authors evaluated two simple classifiers that are based on a learning algorithm (boosting algorithm) 2- CNN classifiers were then trained on ImageNet Large scale	To evaluate deep learning classification techniques in a real maternal- fetal clinical environment to identify planes in obstetric ultrasound	The authors demonstrated similar performance compared to humans when classifying common planes in fetal ultrasound

(Continues)

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			Visual recognition challenge to identify the following planes: Fetal abdomen, brain, femur and thorax and maternal cervical length		
Jiao 2020	45 scans	NR	The deep learning model is based on a self-supervised learning approach to learn meaningful and transferable representations from medical imaging video without any type of human annotation. The model is forced to address anatomy-aware tasks with free supervision from the data itself. The model was designed to correct the order of a reshuffled video clip and at the same time predict the geometric transformation applied to the video clip	To assess the ability of a deep learning model to undergo self-supervised learning using unlabeled data without the need for ground truth annotations from human experts	This self-supervised learning approach can learn the representations and this also applies to standard plane detection and saliency prediction Precision 75.8 ± 1.9 Recall: 76.4 ± 2.7 *F*1 score 75.7 ± 2.0 Best performance compared to other deep learning methods like sono Net
Shozu 2020	256 cases	18–28 weeks GA	The authors used a multiframe method (based on time series data from ultrasound videos) and a cylinder method which uses the shape of the thoracic wall to provide accurate segmentation of the thoracic wall Multiframe+ cylinder method employs predictions from CNN models	To improve segmentation of the thoracic wall which is cylindrical in shape	MFCY (multiframe+ cylinder method) increased the mean values of intersection over union of thoracic wall segmentation from 0.448 to 0.493 for U-Net and from 0.417 to 0.470 for DeepLabv3+ U-Net with MFCY better values of intersection, dice and recall but less precision Intersection: 0.493 Dice: 0.654 Precision: 0.596 Recall: 0.738 DeepLabv3+ with MFCY better values of intersection, dice and recall but less precision Intersection: 0.470 Dice: 0.633 Precision: 0.566 Recall: 0.729
Smail 2020	2420 images from 673 patients	Hydronephrosis on ultrasound	To assess the use of deep convolutional neural networks to grade hydronephrosis on	To prove that a deep learning model is able to depict the grades of hydronephrosis on	94% of the images were classified correctly or within one grade. 51% of the images were

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			ultrasound images according to the society of fetal urology classification system compared to experts	ultrasound at or very close to experts	correctly predicted (*f*1: 0.49) CNN had an average accuracy of 78% and *F*1: 0.78 for low versus high grades accuracy of 71% and *F*1: 0.71 for distinguishing between grade 2 and 3
Xia 2020	7013 images form 1023 pregnant women	Normal pregnancies between 20 and 41 + 6 weeks Class 1: 20–29 + 6 weeks Class 2: 30–36 + 6 weeks Class 3: 37–41 + 6 weeks	A convolutional neural network was established to extract and classify different ultrasound images of the fetal lung in relation to GA. This system was then validated by a 10 fold cross validation	To create a grading model for normal fetal lung gestational age through deep learning mechanisms and to validate this method as well as assess its potential in determining fetal lung maturity	The authors noted the following results. *Sensitivity*: Class 1: 91.7% Class 2: 69.8% Class 3: 86.4% *Specificity* Class 1: 6.8% Class 2: 90% Class 3: 83.1% *Total accuracy*: 83.38% AUC Class 1: 0.982 Class 2: 0.907 Class 3: 0.960
Yang 2020	2081 images for training, 450 for validation, and 727 for testing	NR	The authors used the deep learning models residual U-net and ASPP U-net to improve the accuracy of fetal US segmentation without increasing the depth of the model	To propose a new method to improve the accuracy of fetal US segmentation of CRL, AC and FL	The authors found that the proposed networks could effectively improve the segmentation accuracy of fetal US images. For example, on the AC dataset, the residual U-net performed superior with improvement from 0.8985 to 0.9241 in dice compared with U-net, while ASPP U-net further increases to 0.9412. However, the authors concluded that automatic and accurate segmentation of fetal US images remains challenging due to a variety of interference factors
Burgos-Artizzu 2019	790 fetal lung images	Pregnancies between 24 and 38 + 6 weeks GA in which an ultrasound was obtained within 48 h of delivery, maternal BMI <35, no congenital malformation	An axial section of the fetal thorax at the level of the four-chamber cardiac view is obtained and images were processed using the quantusFLM 3.0, which automatically delineated a ROI in the fetal lung and calculated a NRM risk score based on deep learning techniques	To evaluate the performance of a new version of quantusFLM software for prediction of neonatal respiratory morbidity by ultrasound	QuantusFLM predicted NRM with a sensitivity, specificity, and positive and negative predictive value of 71.0%, 94.7%, 67.9%, and 95.4%, respectively, with an accuracy of 91.5% which is similar to tests based on amniotic fluid and more accurate than gestational age alone

(Continues)

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Chen 2019	2694 2D ultrasound exams	18–22 weeks gestation	Training CNNs are often limited by the availability of an adequate number of labeled images to train a model accurately. A novel self-supervised learning strategy for medical imaging is proposed by the authors to overcome this limitation. Two small patches of an image are randomly selected and swapped. The CNN is then trained to restore the image back to its original version	To evaluate a novel self-supervised learning strategy based on context restoration in order to better exploit unlabeled images to improve the performance of machine learning models	Context restoration improved SonoNet performance compared to the original study by Baumgartner et al. with a precision of 80.6%, recall of 86% and an *F*1-score of 82.8%
Gomez 2019	8 images and 2 fetuses	Healthy fetuses	The authors proposed an image reconstruction framework to combine a large number of overlapping image patches into a fused reconstruction of the object of interest, that is robust to inconsistencies between patches (e.g. motion artifacts) without explicitly modeling them. The authors proposed a new method based on a convolutional variational autoencoder (β-VAE), and compared it to classical manifold embedding techniques: Linear (MultiDimensional scaling) and nonlinear (Laplacian Eigenmaps)	To propose an image reconstruction framework and apply it to whole-fetus US imaging	The authors concluded that the β-VAE method outperformed all other methods in terms of preservation of patch information and overall image quality
Maraci 2017	323 videos	>28 weeks gestation	The authors developed an image analysis framework for linear ultrasound videos to allow less experienced users of the US to identify structures and interpret images particularly related to fetal presentation and the presence or absence of a fetal heartbeat	To develop a framework for characterizing an ultrasound video obtained from a predefined scan protocol for pregnancies	From a total of 129 unseen videos in the test dataset, 41 videos did not contain the skull or abdomen, which are essential for the detection of fetal presentation From the remaining 88 videos, the presentation was correctly identified in 76 videos sweeps (83.4%)

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
					For the detection of the heartbeat an accuracy of 93.1% was achieved
Li 2017	1300 images	Fetal anomaly scans	The data was fed to a downscaled convolution structure. Then decoding was done through learnable convolution kernels. Convolution layers were added between encoding and decoding that then resulted in classification of fetal tissue and amniotic fluid	To determine the effectiveness of deep learning methods on segmentation of the amniotic fluid and fetal tissues based on ultrasound images	The authors demonstrated: 93% global accuracy 67% fetal body accuracy 78% amniotic fluid accuracy
Zhen 2016	1735 images	20–36 weeks	The authors used very deep convolutional networks to represent fine grained details in an US image Very small convolution filters were used to improve the performance of the model	To propose a framework to automatically detect fetal facial standard plane on US images	The authors noted the following results: Accuracy: 96.99% True positive rate: 96.98% False positive rate: 98.49% Precision: 96.98% Recall: 98.99%
Bonet-Carne 2015	144 neonates	28–39 weeks gestation	A computerized method, termed quantusFLM, based on texture analysis and machine learning algorithms was trained to predict neonatal respiratory morbidity risk on fetal lung ultrasound images. QuantusFLM, was then validated blindly in neonates using lung ultrasound images obtained within 48 h of delivery	To develop and assess the performance of quantusFLM for predicting neonatal respiratory morbidity based on quantitative analysis of the fetal lung by ultrasound	Among the 144 neonates, there were 29 (20.1%) cases of neonatal respiratory morbidity. Quantitative texture analysis predicted neonatal respiratory morbidity with a sensitivity, specificity, positive predictive value and negative predictive value of 86.2%, 87.0%, 62.5% and 96.2%, respectively
Lei 2015	1753 images	20–36 weeks gestation	Densely sampled RootSIFT features are extracted and then encoded by FV. The Fisher network with multi-layer design was developed to extract spatial information to boost the classification performance. Finally, automatic recognition of the fetal facial standard planes is implemented by a SVM classifier based on the SDCA algorithm	To propose a new algorithm for the detection of fetal facial standard planes that include axial, coronal, and sagittal planes	The authors demonstrated an accuracy of 93.27% mAP of 99.19%

(Continues)

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
Yaqub 2015	29,858 images from 2256 scans	Anomaly scans (18–22 + 6 weeks)	The authors assessed the use of guided random forests for identification of key fetal anatomy and image categorization. Random forests classifiers were taught to extract regions inside the images where useful structures exist. This method also utilizes a translation and orientation method that captures the region at multiple spatial resolutions. The authors reported Top 1 and Top 2 accuracy for the model. An image is considered in TPtop2 if its class is within the top 2 probabilities of the algorithm	To propose a new learning method to categorize unlabeled fetal ultrasound images	Accuracy of the top 1 probability of the algorithm is 75% compared to top 2 accuracy of 91%
Lei 2014	486 images	20–36 weeks	An automatic algorithm was developed to address the issue of recognition of standard planes in fetal ultrasound. The dense sampling feature transform descriptor (DSIFT) with aggregating vector method (i.e. FV) was explored for feature extraction. The learning and recognition of the planes were implemented by SVM classifier	To propose an automatic algorithm for the recognition of standard planes (i.e. axial, coronal and sagittal planes) in the fetal ultrasound images	High recognition accuracy was demonstrated with mAP >95% for coronal and >84% mAP for axial planes
Ni 2014	1995 images for training 223 videos for testing	18–40 weeks GA	The region of the abdomen was detected by RF classifiers and three component detectors were trained using the RFs to locate the SB, UV and SP. A SVM was then used to analyze the component detectors and give the conclusion of FASP or non-FASP	To develop a model that can properly detect key anatomical structures, handle anatomic variations, and have the ability to exclude regions that might appear similar to the key anatomic structures	The UV results were the best followed by the SB, SP and ROI. This result may be due to the distinctive appearance of the UV. The AUC values for all the classifiers were above 0.98
Cobo 2012	957 images	20–40 weeks	The authors evaluated automatic quantitative ultrasound analysis	To evaluate the reproducibility and feasibility of a	The authors demonstrated a strong correlation with gestational age

TABLE 7 (Continued)

Author year	Number of patients	Inclusion criteria	Description of artificial intelligence	Objective	Results
			software to extract images and quantify lung texture. Feature transformation and a regression model were then used to correlate the extracted image with the gestational age	quantitative ultrasound analysis in detecting fetal lung texture in correlation with gestational age	with a Pearson correlation of 0.97
Feng 2009	1010 fetal images	21–40 weeks	The authors used constrained marginal space learning for the detection of the fetal face, and a boosting profile to refine the image. An automatic carving algorithm was then used to remove everything obstructing the face	To propose a learning-based approach that combines information from 2D and 3D images to detect fetal face on 3D images	The authors demonstrated a high detection accuracy and the system was able to detect the fetal face in 1 s

Abbreviations: 3D, three dimensional; 4D, four dimensional; AC, abdominal circumference; AF, amniotic fluid; AFI, amniotic fluid index; AI, artificial intelligence; ASPP, atrous spatial pyramid pooling; AUC, area under the curve; CNN, convolutional neural network; CRL, crown rump length; DSC, dice similarity coefficient; FASP, fetal abdominal standard plane; FL, femur length; FV, Fisher vector; GA, gestational age; GDM, gestational diabetes mellitus; KNN, K-nearest neighbours; mAP, mean average precision; ML, machine learning; NR, not reported; OA, occipital anterior; PE, preeclampsia; PPROM, preterm premature rupture of membranes; ROI, region of interest; RootSIFT, root scale invariant feature transform; SB, stomach bubble; SDCA, stochastic dual coordinate ascent; SMOTE, synthetic minority oversampling technique; SP, spine; SVM, support vector machine; UCI, umbilical coiling index; US, ultrasound; UV, umbilical vein.

networks to automatically detect various cardiac views and distinguish between normal and abnormal on fetal echo images, with the authors reporting an AUC of 0.99, a sensitivity 95% and a specificity of 96%.[77] In addition to shortening the learning curve for sonographers and physicians to become proficient at ultrasound, AI also has the potential to improve efficiency by shortening scanning time. Matthews et al. evaluated an AI system that provided feedback to the sonographer during live scanning via a traffic light system to inform the sonographer that the required anatomic views had been obtained in addition to automatically measuring fetal biometry.[105]

Our review also highlights the potential of AI in low resource settings. Several studies focused on AI detection of placental location and categorization of the placenta as normal, low-lying, or placenta previa.[27,28] Arroyo et al. assessed the use of AI for fetal presentation, placental location and fetal biometry for estimation of gestational age using standardized ultrasound sweeps obtained in the absence of trained sonographers.[33] Timely identification and transfer of patients with placenta previa and other placental disorders have the potential to significantly improve patient outcomes as rates of up to 50% perinatal mortality have been described amongst patients with antepartum bleeding in low income countries.[139] Accurate assessment of gestational age in resource limited settings also allows for the potential transfer to facilities with higher levels of neonatal care for preterm patients at risk of delivery.

AI may allow new uses of ultrasound in obstetrics, beyond the capabilities of human experts. Nambuerete et al. developed an AI system to predict gestational age and neurodevelopmental maturation of the fetus from 3D images,[94] whereas Burgos-Artizzu et al. assessed the automated analysis of fetal brain morphology on standard cranial ultrasound sections, to estimate the gestational age.[58] It remains to be seen if AI will lead to new methods of estimating gestational age with increased accuracy compared with standard fetal biometry. Several studies focused on fetal lung ultrasound for assessment of fetal lung maturity and prediction of neonatal respiratory morbidity.[112,128] Recent literature has suggested the need for a more judicious approach to administration of antenatal corticosteroids due to growing concerns about potential long-term neurodevelopmental effects of in-utero corticosteroid exposure.[140,141] Perhaps AI will be able to predict those who will benefit most from antenatal corticosteroids while minimizing unnecessary exposure. AI in obstetric ultrasound may be useful to predict adverse pregnancy outcomes, thus identifying patients requiring more intensive surveillance. Looney et al. assessed automated placental volume in the first trimester for the prediction of SGA neonates at birth and Gupta et al. compared placental quantitative image texture throughout pregnancy in patients with hypertensive disorders to controls with normal pregnancy outcomes, demonstrating promising results for the prediction of hypertensive disorders.[24,31]

TABLE 8 Summary of studies examining the use of artificial intelligence in obstetric ultrasound.

Study type	Number of studies
First trimester ultrasound	11
Measurement of nuchal translucency	3
Detection of gestational sac	3
Detection of the mid-sagittal plane	3
Fetal biometry	1
Measurement of cerebral cortex	1
Placenta ultrasound	8
Segmentation of the placenta only	2
First trimester placenta volume for the prediction of SGA neonates	1
Categorization of placental location	1
Detection of placental location	1
Staging of placental maturity	1
Detection of lacunae	1
Comparison of placental texture throughout pregnancy in patients with hypertensive disorders to normotensive patients	1
Fetal biometry	47
Measurement of fetal biometry	10
Fetal head measurements	20
Assessment of fetal brain morphology for estimation of gestational age	1
Femur length only	5
Abdominal circumference only	2
Detection of fetal abdominal standard plane	4
Various combinations of ≥2 fetal biometry measurements	5
Fetal cardiac imaging	10
Detection of substructures of the heart to calculate abnormality score	1
Detection of outflow tracts	1
Detection of standardized fetal heart views and characterization as normal or abnormal	1
Assessment of the ventricular septum	1
Detection of four chamber view	3
Quality control of the four chamber view	1
Automatic segmentation of the fetal heart and lungs	1
Fetal neurosonography	20
Localization of planes in the fetal brain using 3D volumes	3
Segmentation of intracranial structures or fetal cranium	5
Detection of intracranial planes using 2D images	2
Assessment of neurodevelopmental maturation	1
Automated detection and measurement of ≥1 intracranial structure	5
Detection of specified intracranial abnormalities	3
Classification of fetal head biometry as normal or abnormal	1
Anatomical evaluation of the fetus	6
Detection of standard anatomic planes	4
Classification of fetal structures	1
Assessment of intra and inter observer variability during morphology ultrasound	1
Other	25
Assessment of fetal lungs	4
Assessment of the ability of deep learning models to undergo self-supervised learning using fetal images/video	3
Detection of fetal facial planes	2
Assessment of amniotic fluid	2
Other	13
Total	127

Abbreviations: 2D, two dimensional; 3D, three dimensional; SGA, small for gestational age.

Our review has many strengths. To our knowledge, this is the first scoping literature review of AI in obstetric ultrasound to date. We have synthesized all published studies examining the use of AI in obstetric ultrasound, providing a comprehensive overview of the current capabilities, challenges and potential future uses of AI in obstetric ultrasound. There are several limitations to this review. Firstly, the heterogeneity of studies included in this review does not allow meaningful comparison regarding which AI methods are superior for obstetric ultrasound. Furthermore, the majority of included studies are retrospective in nature with very few prospective studies to date examining AI in obstetric ultrasound. Therefore, it remains unknown if AI in obstetric ultrasound will have the ability to improve maternal and fetal outcomes in a real-life setting.

The use of AI in obstetric ultrasound will likely increase in the coming decades. Many of the latest ultrasound systems have integrated intelligent applications, often to obtain measurements based on standard plane detection of an image obtained by the sonographer. AI holds the potential to improve the ultrasound efficiency, decrease interobserver variability, improve detection of congenital malformations and shorten the training duration to become proficient at obstetric ultrasound. Future research should focus on determining the optimal AI techniques in obstetric ultrasound and assessing if maternal and fetal outcomes are improved with the use of AI in obstetric ultrasound.

ACKNOWLEDGMENTS

None.

CONFLICT OF INTEREST STATEMENT

The authors did not report any potential conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Rebecca Horgan https://orcid.org/0000-0002-2752-4061

REFERENCES

1. Artificial intelligence technologies. United Kingdom Engineering and Physical Sciences Research Council. Accessed December 12, 2022. https://epsrc.ukri.org/research/ourportfolio/researchareas/ait/
2. He F, Wang Y, Xiu Y, sinclair Y, Chen L. Artificial intelligence in prenatal ultrasound diagnosis. *Front Med*. 2021;8:729978. https://doi.org/10.3389/fmed.2021.729978
3. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin*. 2019;69(2):127-157. https://doi.org/10.3322/caac.21552
4. Choi YJ, Baek JH, Park HS, et al. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment. *Thyroid*. 2017;27(4):546-552. https://doi.org/10.1089/thy.2016.0372
5. Nishida N, Kudo M. Artificial intelligence in medical imaging and its application in sonography for the management of liver tumor. *Front Oncol*. 2020;10:594580. https://doi.org/10.3389/fonc.2020.594580
6. Liang X, Yu J, Liao J, Chen Z. Convolutional neural network for breast and thyroid nodules diagnosis in ultrasound imaging. *BioMed Res Int*. 2020;2020:1763803-1763809. https://doi.org/10.1155/2020/1763803
7. Lei YM, Yin M, Yu MH, et al. Artificial intelligence in medical imaging of the breast. *Front Oncol*. 2021;11:600557. https://doi.org/10.3389/fonc.2021.600557
8. Zhang L, Ye X, Lambrou T, Duan W, Allinson N, Dudley NJ. A supervised texton based approach for automatic segmentation and measurement of the fetal head and femur in 2D ultrasound images. *Phys Med Biol*. 2016;61(3):1095-1115. https://doi.org/10.1088/0031-9155/61/3/1095
9. Li P, Zhao H, Liu P, Cao F. Automated measurement network for accurate segmentation and parameter modification in fetal head ultrasound images. *Med Biol Eng Comput*. 2020;58(11):2879-2892. https://doi.org/10.1007/s11517-020-02242-5
10. van den Heuvel TLA, Petros H, Santini S, de Korte CL, van Ginneken B. Automated fetal head detection and circumference estimation from free-hand ultrasound sweeps using deep learning in resource-limited countries. *Ultrasound Med Biol*. 2019;45(3):773-785. https://doi.org/10.1016/j.ultrasmedbio.2018.09.015
11. Sinclair M, Baumgartner CF, Matthew J, et al. *Human-Level Performance on Automatic Head Biometrics in Fetal Ultrasound Using Fully Convolutional Neural Networks*. IEEE; 2018:714-717.
12. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473. https://doi.org/10.7326/m18-0850
13. Sciortino G, Tegolo D, Valenti C. Automatic detection and measurement of nuchal translucency. *Comput Biol Med*. 2017;82:12-20. https://doi.org/10.1016/j.compbiomed.2017.01.008
14. Park J, Sofka M, Lee S, Kim D, Zhou SK. Automatic nuchal translucency measurement from ultrasonography. *Med Image Comput Comput Assist Interv*. 2013;16(Pt 3):243-250. https://doi.org/10.1007/978-3-642-40760-4_31
15. Deng Y, Wang Y, Chen P, Yu J. A hierarchical model for automatic nuchal translucency detection from ultrasound images. *Comput Biol Med*. 2012;42(6):706-713. https://doi.org/10.1016/j.compbiomed.2012.04.002
16. Borenstein M, Azumendi Perez G, Molina Garcia F, Romero M, Anderica JR. Gestational sac volume: comparison between SonoAVC and VOCAL measurements at 11 + 0 to 13 + 6 weeks of gestation. *Ultrasound Obstet Gynecol*. 2009;34(5):510-514. https://doi.org/10.1002/uog.7342
17. Zhang L, Chen S, Chin CT, Wang T, Li S. Intelligent scanning: automated standard plane selection and biometric measurement of early gestational sac in routine ultrasound examination. *Med Phys*. 2012;39(8):5015-5027. https://doi.org/10.1118/1.4736415
18. Zhang L, Chen S, Li S, Wang T. Automatic measurement of early gestational sac diameters from one scan session. 2011.
19. Sciortino G, Orlandi E, Valenti C, Tegolo D. Wavelet analysis and neural network classifiers to detect mid-sagittal sections for nuchal translucency measurement. *Image Anal Stereol*. 2016;35(2):105. https://doi.org/10.5566/ias.1352
20. Tsai PY, Hung CH, Chen CY, Sun YN. Automatic fetal middle sagittal plane detection in ultrasound using generative adversarial network. *Diagnostics*. 2020;11(1):21. https://doi.org/10.3390/diagnostics11010021
21. Nie S, Yu J, Chen P, Wang Y, Zhang JQ. Automatic detection of standard sagittal plane in the first trimester of pregnancy using 3-D ultrasound data. *Ultrasound Med Biol*. 2017;43(1):286-300. https://doi.org/10.1016/j.ultrasmedbio.2016.08.034
22. Ryou H, Yaqub M, Cavallaro A, Papageorghiou AT, Alison Noble J. Automated 3D ultrasound image analysis for first trimester assessment of fetal health. *Phys Med Biol*. 2019;64(18):185010. https://doi.org/10.1088/1361-6560/ab3ad1
23. Gofer S, Haik O, Bardin R, Gilboa Y, Perlman S. Machine learning algorithms for classification of first-trimester fetal brain ultrasound images. *J Ultrasound Med*. 2022;41(7):1773-1779. https://doi.org/10.1002/jum.15860
24. Looney P, Stevenson GN, Nicolaides KH, et al. Fully automated, real-time 3D ultrasound segmentation to estimate first trimester placental volume using deep learning. *JCI Insight*. 2018;3(11). https://doi.org/10.1172/jci.insight.120178
25. Torrents-Barrena J, Monill N, Piella G, et al. Assessment of radiomics and deep learning for the segmentation of fetal and maternal anatomy in magnetic resonance imaging and ultrasound. *Acad Radiol*. 2021;28(2):173-188. https://doi.org/10.1016/j.acra.2019.11.006
26. Hu R, Singla R, Yan R, Mayer C, Rohling RN. Automated placenta segmentation with a convolutional neural network weighted by acoustic shadow detection. 2019:6718-6723.
27. Schilpzand M, Neff C, van Dillen J, et al. Automatic placenta localization from ultrasound imaging in a resource-limited setting using a predefined ultrasound acquisition protocol and deep learning. *Ultrasound Med Biol*. 2022;48(4):663-674. https://doi.org/10.1016/j.ultrasmedbio.2021.12.006
28. Saavedra AC, Arroyo J, Tamayo L, Egoavil M, Ramos B, Castaneda B. Automatic ultrasound assessment of placenta previa during the third trimester for rural areas. 2020:1-4.
29. Lei B, Yao Y, Chen S, et al. Discriminative learning for automatic staging of placental maturity via multi-layer Fisher vector. *Sci Rep*. 2015;5(1):12818. https://doi.org/10.1038/srep12818
30. Qi H, Collins S, Noble JA. Automatic lacunae localization in placental ultrasound images via layer aggregation. *Med Image Comput Comput Assist Interv*. 2018;11071:921-929. https://doi.org/10.1007/978-3-030-00934-2_102

31. Gupta K, Balyan K, Lamba B, Puri M, Sengupta D, Kumar M. Ultrasound placental image texture analysis using artificial intelligence to predict hypertension in pregnancy. *J Matern Fetal Neonatal Med*. 2022;35(25):5587-5594. https://doi.org/10.1080/14767058.2021.1887847
32. Plotka S, Klasa A, Lisowska A, et al. Deep learning fetal ultrasound video model match human observers in biometric measurements. *Phys Med Biol*. 2022;16(4):67. https://doi.org/10.1088/1361-6560/ac4d85
33. Arroyo J, Marini TJ, Saavedra AC, et al. No sonographer, no radiologist: new system for automatic prenatal detection of fetal biometry, fetal presentation, and placental location. *PLoS One*. 2022; 17(2):e0262107. https://doi.org/10.1371/journal.pone.0262107
34. Prieto JC, Shah H, Rosenbaum AJ, et al. An automated framework for image classification and segmentation of fetal ultrasound images for gestational age estimation. *Proc SPIE-Int Soc Opt Eng*. 2021;11596. https://doi.org/10.1117/12.2582243
35. Ghelich Oghli M, Shabanzadeh A, Moradi S, et al. Automatic fetal biometry prediction using a novel deep convolutional network architecture. *Phys Med*. 2021;88:127-137. https://doi.org/10.1016/j.ejmp.2021.06.020
36. Carneiro G, Georgescu B, Good S, Comaniciu D. Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. *IEEE Trans Med Imag*. 2008;27(9):1342-1355. https://doi.org/10.1109/TMI.2008.928917
37. Carneiro G, Georgescu B, Good S, Comaniciu D. Automatic fetal measurements in ultrasound using constrained probabilistic boosting tree. *Med Image Comput Comput Assist Interv*. 2007;10(Pt 2):571-579. https://doi.org/10.1007/978-3-540-75759-7_69
38. Rueda S, Fathima S, Knight CL, et al. Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge. *IEEE Trans Med Imag*. 2014;33(4):797-813. https://doi.org/10.1109/TMI.2013.2276943
39. Bano S, Dromey B, Vasconcelos F, et al. *AutoFB: Automating Fetal Biometry Estimation from Standard Ultrasound Planes*. Springer International Publishing; 2021:228-238.
40. Luo D, Wen H, Peng G, et al. A prenatal ultrasound scanning approach: one-touch technique in second and third trimesters. *Ultrasound Med Biol*. 2021;47(8):2258-2265. https://doi.org/10.1016/j.ultrasmedbio.2021.04.020
41. Rong Y, Xiang D, Zhu W, et al. Deriving external forces via convolutional neural networks for biomedical image segmentation. *Biomed Opt Express*. 2019;10(8):3800-3814. https://doi.org/10.1364/BOE.10.003800
42. Ambroise Grandjean G, Hossu G, Bertholdt C, Noble P, Morel O, Grange G. Artificial intelligence assistance for fetal head biometry: assessment of automated measurement software. *Diagn Interv Imaging*. 2018;99(11):709-716. https://doi.org/10.1016/j.diii.2018.08.001
43. Pluym ID, Afshar Y, Holliman K, et al. Accuracy of automated three-dimensional ultrasound imaging technique for fetal head biometry. *Ultrasound Obstet Gynecol*. 2021;57(5):798-803. https://doi.org/10.1002/uog.22171
44. Kim HP, Lee SM, Kwon JY, Park Y, Kim KC, Seo JK. Automatic evaluation of fetal head biometry from ultrasound images using machine learning. *Physiol Meas*. 2019;40(6):065009. https://doi.org/10.1088/1361-6579/ab21ac
45. Perez-Gonzalez JL, Muńoz JCB, Porras MCR, Arámbula-Cosío F, Medina-Bańuelos V. Automatic fetal head measurements from ultrasound images using optimal ellipse detection and texture maps. In: *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014*. 2015:329-332. Chapter 85. *IFMBE Proceedings*.
46. Li J, Wang Y, Lei B, et al. Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting. *IEEE J Biomed Health Inform*. 2018;22(1):215-223. https://doi.org/10.1109/JBHI.2017.2703890
47. Foi A, Maggioni M, Pepe A, et al. Difference of Gaussians revolved along elliptical paths for ultrasound fetal head segmentation. *Comput Med Imaging Graph*. 2014;38(8):774-784. https://doi.org/10.1016/j.compmedimag.2014.09.006
48. Al-Bander B, Alzahrani T, Alzahrani S, Williams BM, Zheng Y. *Improving Fetal Head Contour Detection by Object Localisation with Deep Learning*. Springer International Publishing; 2020:142-150.
49. Fiorentino MC, Moccia S, Capparuccini M, Giamberini S, Frontoni E. A regression framework to head-circumference delineation from US fetal images. *Comput Methods Progr Biomed*. 2021;198:105771. https://doi.org/10.1016/j.cmpb.2020.105771
50. Moccia S, Fiorentino MC, Frontoni E. Mask-R[Formula: see text] CNN: a distance-field regression version of Mask-RCNN for fetal-head delineation in ultrasound images. *Int J Comput Assist Radiol Surg*. 2021;16(10):1711-1718. https://doi.org/10.1007/s11548-021-02430-0
51. Wang X, Wang W, Cai X. Automatic measurement of fetal head circumference using a novel GCN-assisted deep convolutional network. *Comput Biol Med*. 2022;145:105515. https://doi.org/10.1016/j.compbiomed.2022.105515
52. Zeng Y, Tsui PH, Wu W, Zhou Z, Wu S. Fetal ultrasound image segmentation for automatic head circumference biometry using deeply supervised attention-gated V-Net. *J Digit Imag*. 2021;34(1):134-148. https://doi.org/10.1007/s10278-020-00410-5
53. Zhang J, Petitjean C, Ainouz S. Segmentation-based vs. regression-based biomarker estimation: a case study of fetus head circumference assessment from ultrasound images. *J Imaging*. 2022;8(2):8. https://doi.org/10.3390/jimaging8020023
54. Sobhaninia Z, Rafiei S, Emami A, et al. Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning. *IEEE*. 2019:6545-6548.
55. Zhang J, Petitjean C, Lopez P, Ainouz S. Direct estimation of fetal head circumference from ultrasound images based on regression CNN. In: *Presented at: Proceedings of the Third Conference on Medical Imaging with Deep Learning*; 2020. Proceedings of Machine Learning Research. Accessed December 12, 2022. https://proceedings.mlr.press/v121/zhang20a.html
56. Sun C. Automatic fetal head measurements from ultrasound images using circular shortest paths. n.d..
57. Rajinikanth V, Dey N, Kumar R, Panneerselvam J, Raja NSM. Fetal head periphery extraction from ultrasound image using Jaya algorithm and Chan-Vese segmentation. *Procedia Comput Sci*. 2019/01/01/ 2019;152:66-73. https://doi.org/10.1016/j.procs.2019.05.028
58. Burgos-Artizzu XP, Coronado-Gutierrez D, Valenzuela-Alcaraz B, et al. Analysis of maturation features in fetal brain ultrasound via artificial intelligence for the estimation of gestational age. *Am J Obstet Gynecol MFM*. 2021;3(6):100462. https://doi.org/10.1016/j.ajogmf.2021.100462
59. Zhu F, Liu M, Wang F, Qiu D, Li R, Dai C. Automatic measurement of fetal femur length in ultrasound images: a comparison of random forest regression model and SegNet. *Math Biosci Eng*. 2021;18(6):7790-7805. https://doi.org/10.3934/mbe.2021387
60. Hur H, Kim YH, Cho HY, et al. Feasibility of three-dimensional reconstruction and automated measurement of fetal long bones using 5D Long Bone. *Obstet Gynecol Sci*. 2015;58(4):268-276. https://doi.org/10.5468/ogs.2015.58.4.268
61. Wang CW. Automatic entropy-based femur segmentation and fast length measurement for fetal ultrasound images. 2014:1-5.
62. Shrimali V, Anand RS, Kumar V. Improved segmentation of ultrasound images for fetal biometry, using morphological operators. 2009:459-462.
63. Mukherjee P, Swamy G, Gupta M, Patil U, Krishnan KB. Automatic detection and measurement of femur length from fetal

ultrasonography. In: *Presented at: Medical Imaging 2010: Ultrasonic Imaging, Tomography, and Therapy*. 2010.
64. Jang J, Park Y, Kim B, Lee SM, Kwon JY, Seo JK. Automatic estimation of fetal abdominal circumference from ultrasound images. *IEEE J Biomed Health Inform*. 2018;22(5):1512-1520. https://doi.org/10.1109/JBHI.2017.2776116
65. Kim B, Kim KC, Park Y, Kwon JY, Jang J, Seo JK. Machine-learning-based automatic identification of fetal abdominal circumference from ultrasound images. *Physiol Meas*. 2018;39(10):105007. https://doi.org/10.1088/1361-6579/aae255
66. Chen H, Ni D, Qin J, et al. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. *IEEE J Biomed Health Inform*. 2015;19(5):1627-1636. https://doi.org/10.1109/JBHI.2015.2425041
67. Chen H, Ni D, Yang X, Li S, Heng PA. *Fetal Abdominal Standard Plane Localization through Representation Learning with Knowledge Transfer*. Springer International Publishing; 2014:125-132.
68. Rahmatullah B, Sarris I, Papageorghiou A, Noble JA. Quality control of fetal ultrasound images: detection of abdomen anatomical landmarks using AdaBoost. 2011:6-9.
69. Wu L, Cheng JZ, Li S, Lei B, Wang T, Ni D. FUIQA: fetal ultrasound image quality assessment with deep convolutional networks. *IEEE Trans Cybern*. 2017;47(5):1336-1349. https://doi.org/10.1109/TCYB.2017.2671898
70. Ashkani Chenarlogh V, Ghelich Oghli M, Shabanzadeh A, et al. Fast and accurate U-Net model for fetal ultrasound image segmentation. *Ultrason Imag*. 2022;44(1):25-38. https://doi.org/10.1177/01617346211069882
71. Salim I, Cavallaro A, Ciofolo-Veit C, et al. Evaluation of automated tool for two-dimensional fetal biometry. *Ultrasound Obstet Gynecol*. 2019;54(5):650-654. https://doi.org/10.1002/uog.20185
72. Wu L, Xin Y, Li S, Wang T, Heng P-A, Ni D. Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation. In: *Presented at: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. 2017.
73. Ponomarev GV, Gelfand MS, Kazanov MD. A multilevel thresholding combined with edge detection and shape-based recognition for segmentation of fetal ultrasound images. In: *Proceedings of Challenge US: Biometric Measurements from Fetal Ultrasound Images, ISBI*. 2012:17-19.
74. Miyagi Y, Miyake T. Potential of artificial intelligence for estimating Japanese fetal weights. *Acta Med Okayama*. 2020;74(6):483-493. https://doi.org/10.18926/amo/61207
75. Sakai A, Komatsu M, Komatsu R, et al. Medical professional enhancement using explainable artificial intelligence in fetal cardiac ultrasound screening. *Biomedicines*. 2022;10(3):10. https://doi.org/10.3390/biomedicines10030551
76. Abuhamad A, Falkensammer P, Reichartseder F, Zhao Y. Automated retrieval of standard diagnostic fetal cardiac ultrasound planes in the second trimester of pregnancy: a prospective evaluation of software. *Ultrasound Obstet Gynecol*. 2008;31(1):30-36. https://doi.org/10.1002/uog.5228
77. Arnaout R, Curran L, Zhao Y, Levine JC, Chinn E, Moon-Grady AJ. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. *Nat Med*. 2021;27(5):882-891. https://doi.org/10.1038/s41591-021-01342-5
78. Herling L, Johnson J, Ferm-Widlund K, Zamprakou A, Westgren M, Acharya G. Automated quantitative evaluation of fetal atrioventricular annular plane systolic excursion. *Ultrasound Obstet Gynecol*. 2021;58(6):853-863. https://doi.org/10.1002/uog.23703
79. Dozen A, Komatsu M, Sakai A, et al. Image segmentation of the ventricular septum in fetal cardiac ultrasound videos based on deep learning using time-series information. *Biomolecules*. 2020;10(11):10. https://doi.org/10.3390/biom10111526
80. Qiao S, Pan S, Luo G, et al. A pseudo-siamese feature fusion generative adversarial network for synthesizing high-quality fetal four-chamber views. *IEEE J Biomed Health Inform*. 2022;27(3):1193-1204. https://doi.org/10.1109/JBHI.2022.3143319
81. Qiao S, Pang S, Luo G, Pan S, Chen T, Lv Z. FLDS: an intelligent feature learning detection system for visualizing medical images supporting fetal four-chamber views. *IEEE J Biomed Health Inform*. 2017;26(10):4814-4825. https://doi.org/10.1109/JBHI.2021.3091579
82. Dong J, Liu S, Liao Y, et al. A generic quality control framework for fetal ultrasound cardiac four-chamber planes. *IEEE J Biomed Health Inform*. 2020;24(4):931-942. https://doi.org/10.1109/JBHI.2019.2948316
83. Sundaresan V, Bridge CP, Ioannou C, Noble JA. Automated Characterization of the Fetal Heart in Ultrasound Images Using Fully Convolutional Neural Networks. 2017:671-674.
84. Xi J, Chen J, Wang Z, et al. Simultaneous segmentation of fetal hearts and lungs for medical ultrasound images via an efficient multi-scale model integrated with attention mechanism. *Ultrason Imag*. 2021;43(6):308-319. https://doi.org/10.1177/01617346211042526
85. Di Vece C, Dromey B, Vasconcelos F, David AL, Peebles D, Stoyanov D. Deep learning-based plane pose regression in obstetric ultrasound. *Int J Comput Assist Radiol Surg*. 2022;17(5):833-839. https://doi.org/10.1007/s11548-022-02609-z
86. Namburete AIL, Xie W, Yaqub M, Zisserman A, Noble JA. Fully-automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning. *Med Image Anal*. 2018;46:1-14. https://doi.org/10.1016/j.media.2018.02.006
87. Skelton E, Matthew J, Li Y, et al. Towards automated extraction of 2D standard fetal head planes from 3D ultrasound acquisitions: a clinical evaluation and quality assessment comparison. *Radiography*. 2021;27(2):519-526. https://doi.org/10.1016/j.radi.2020.11.006
88. Hesse LS, Aliasi M, Moser F, et al. Subcortical segmentation of the fetal brain in 3D ultrasound using deep learning. *Neuroimage*. 2022;254:119117. https://doi.org/10.1016/j.neuroimage.2022.119117
89. Huang R, Xie W, Alison Noble J. VP-Nets: efficient automatic localization of key brain structures in 3D fetal neurosonography. *Med Image Anal*. 2018;47:127-139. https://doi.org/10.1016/j.media.2018.04.004
90. Yaqub M, Kelly B, Papageorghiou AT, Noble JA. A deep learning solution for automatic fetal neurosonographic diagnostic plane verification using clinical standard constraints. *Ultrasound Med Biol*. 2017;43(12):2925-2933. https://doi.org/10.1016/j.ultrasmedbio.2017.07.013
91. Sofka M, Zhang J, Good S, Zhou SK, Comaniciu D. Automatic detection and measurement of structures in fetal head ultrasound volumes using sequential estimation and integrated detection network (IDN). *IEEE Trans Med Imag*. 2014;33(5):1054-1070. https://doi.org/10.1109/TMI.2014.2301936
92. Montero A, Bonet-Carne E, Burgos-Artizzu XP. Generative adversarial networks to improve fetal brain fine-grained plane classification. *Sensors*. 2021;29(23):21. https://doi.org/10.3390/s21237975
93. Qu R, Xu G, Ding C, Jia W, Sun M. Standard plane identification in fetal brain ultrasound scans using a differential convolutional neural network. *IEEE Access*. 2020;8:83821-83830. https://doi.org/10.1109/access.2020.2991845
94. Namburete AI, Stebbing RV, Kemp B, Yaqub M, Papageorghiou AT, Alison Noble J. Learning-based prediction of gestational age from ultrasound images of the fetal brain. *Med Image Anal*. 2015;21(1):72-86. https://doi.org/10.1016/j.media.2014.12.006
95. Alansary A, Oktay O, Li Y, et al. Evaluating reinforcement learning agents for anatomical landmark detection. *Med Image Anal*. 2019;53:156-164. https://doi.org/10.1016/j.media.2019.02.007

96. Chen X, He M, Dan T, et al. Automatic measurements of fetal lateral ventricles in 2D ultrasound images using deep learning. *Front Neurol*. 2020;11:526. https://doi.org/10.3389/fneur.2020.00526
97. Huang R, Namburete A, Noble A. Learning to segment key clinical anatomical structures in fetal neurosonography informed by a region-based descriptor. *J Med Imaging*. 2018;5(1):014007. https://doi.org/10.1117/1.JMI.5.1.014007
98. Bullard KA, Shaffer BL, Greiner KS, Skeith AE, Rodriguez MI, Caughey AB. Twenty-week abortion bans on pregnancies with a congenital diaphragmatic hernia: a cost-effectiveness analysis. *Obstet Gynecol*. 2018;131(3):581-590. https://doi.org/10.1097/AOG.0000000000002483
99. Yaqub M, Napolitano R, Ioannou C, Papageorghiou AT, Noble JA. Automatic detection of local fetal brain structures in ultrasound images. In: *Presented at: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. 2012.
100. Lin M, He X, Guo H, et al. Use of real-time artificial intelligence in detection of abnormal image patterns in standard sonographic reference planes in screening for fetal intracranial malformations. *Ultrasound Obstet Gynecol*. 2022;59(3):304-316. https://doi.org/10.1002/uog.24843
101. Xie B, Lei T, Wang N, et al. Computer-aided diagnosis for fetal brain ultrasound images using deep convolutional neural networks. *Int J Comput Assist Radiol Surg*. 2020;15(8):1303-1312. https://doi.org/10.1007/s11548-020-02182-3
102. Sahli H, Mouelhi A, Ben Slama A, Sayadi M, Rachdi R. Supervised classification approach of biometric measures for automatic fetal defect screening in head ultrasound images. *J Med Eng Technol*. 2019;43(5):279-286. https://doi.org/10.1080/03091902.2019.1653389
103. Namburete AIL, Noble JA. Fetal cranial segmentation in 2D ultrasound images using shape properties of pixel clusters. 2013:720-723.
104. Xie HN, Wang N, He M, et al. Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. *Ultrasound Obstet Gynecol*. 2020;56(4):579-587. https://doi.org/10.1002/uog.21967
105. Matthew J, Skelton E, Day TG, et al. Exploring a new paradigm for the fetal anomaly ultrasound scan: artificial intelligence in real time. *Prenat Diagn*. 2022;42(1):49-59. https://doi.org/10.1002/pd.6059
106. Baumgartner CF, Kamnitsas K, Matthew J, et al. SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Trans Med Imag*. 2017;36(11):2204-2215. https://doi.org/10.1109/TMI.2017.2712367
107. Sharma H, Droste R, Chatelain P, Drukker L, Papageorghiou AT, Noble JA. Spatio-temporal partitioning and description of full-length routine fetal anomaly ultrasound scans. *Proc IEEE Int Symp Biomed Imaging*. 2019;16:987-990. https://doi.org/10.1109/ISBI.2019.8759149
108. Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images. *Med Image Anal*. 2019;53:197-207. https://doi.org/10.1016/j.media.2019.01.012
109. Sridar P, Kumar A, Quinton A, Nanan R, Kim J, Krishnakumar R. Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks. *Ultrasound Med Biol*. 2019;45(5):1259-1273. https://doi.org/10.1016/j.ultrasmedbio.2018.11.016
110. Chen H, Wu L, Dou Q, et al. Ultrasound standard plane detection using a composite neural network framework. *IEEE Trans Cybern*. 2017;47(6):1576-1586. https://doi.org/10.1109/TCYB.2017.2685080
111. Burgos-Artizzu XP, Perez-Moreno A, Coronado-Gutierrez D, Gratacos E, Palacio M. Evaluation of an improved tool for non-invasive prediction of neonatal respiratory morbidity based on fully automated fetal lung ultrasound analysis. *Sci Rep*. 2019;9(1):1950. https://doi.org/10.1038/s41598-019-38576-w
112. Bonet-Carne E, Palacio M, Cobo T, et al. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. *Ultrasound Obstet Gynecol*. 2015;45(4):427-433. https://doi.org/10.1002/uog.13441
113. Cobo T, Bonet-Carne E, Martinez-Terron M, et al. Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. *Fetal Diagn Ther*. 2012;31(4):230-236. https://doi.org/10.1159/000335349
114. Xia TH, Tan M, Li JH, Wang JJ, Wu QQ, Kong DX. Establish a normal fetal lung gestational age grading model and explore the potential value of deep learning algorithms in fetal lung maturity evaluation. *Chin Med J (Engl)*. 2021;134(15):1828-1837. https://doi.org/10.1097/CM9.0000000000001547
115. Jiao J, Droste R, Drukker L, Papageorghiou AT, Noble JA. Self-supervised representation learning for ultrasound video. *Proc IEEE Int Symp Biomed Imaging*. 2020;2020:1847-1850. https://doi.org/10.1109/ISBI45749.2020.9098666
116. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. Self-supervised learning for medical image analysis using image context restoration. *Med Image Anal*. 2019;58:101539. https://doi.org/10.1016/j.media.2019.101539
117. Yaqub M, Kelly B, Papageorghiou AT, Noble JA. Guided random forests for identification of key fetal anatomy and image categorization in ultrasound scans. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015:687-694. Chapter 82. *Lecture Notes in Computer Science*.
118. Lei B, Tan EL, Chen S, et al. Automatic recognition of fetal facial standard plane in ultrasound image via fisher vector. *PLoS One*. 2015;10(5):e0121838. https://doi.org/10.1371/journal.pone.0121838
119. Zhen Y, Dong N, Siping C, Shengli L, Tianfu W, Baiying L. Fetal facial standard plane recognition via very deep convolutional networks. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016;2016:627-630. https://doi.org/10.1109/embc.2016.7590780
120. Cho HC, Sun S, Min Hyun C, et al. Automated ultrasound assessment of amniotic fluid index using deep learning. *Med Image Anal*. 2021;69:101951. https://doi.org/10.1016/j.media.2020.101951
121. Li Y, Xu R, Ohya J, Iwata H. Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers. 2017:1485-1488.
122. Ghi T, Conversano F, Ramirez Zegarra R, et al. Novel artificial intelligence approach for automatic differentiation of fetal occiput anterior and non-occiput anterior positions during labor. *Ultrasound Obstet Gynecol*. 2022;59(1):93-99. https://doi.org/10.1002/uog.23739
123. Du Y, Fang Z, Jiao J, et al. Application of ultrasound-based radiomics technology in fetal-lung-texture analysis in pregnancies complicated by gestational diabetes and/or pre-eclampsia. *Ultrasound Obstet Gynecol*. 2021;57(5):804-812. https://doi.org/10.1002/uog.22037
124. Miyagi Y, Hata T, Bouno S, Koyanagi A, Miyake T. Recognition of facial expression of fetuses by artificial intelligence (AI). *J Perinat Med*. 2021;49(5):596-603. https://doi.org/10.1515/jpm-2020-0537
125. Ni D, Yang X, Chen X, et al. Standard plane localization in ultrasound by radial component model and selective search. *Ultrasound Med Biol*. 2014;40(11):2728-2742. https://doi.org/10.1016/j.ultrasmedbio.2014.06.006

126. Feng S, Zhou SK, Good S, Comaniciu D. Automatic fetal face detection from ultrasound volumes via learning 3D and 2D information. 2009:2488-2495.
127. Maraci MA, Bridge CP, Napolitano R, Papageorghiou A, Noble JA. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. *Med Image Anal*. 2017;37:22-36. https://doi.org/10.1016/j.media.2017.01.003
128. Shozu K, Komatsu M, Sakai A, et al. Model-agnostic method for thoracic wall segmentation in fetal ultrasound videos. *Biomolecules*. 2020;10(12):10. https://doi.org/10.3390/biom10121691
129. Pradipta GA, Wardoyo R, Musdholifah A, Sanjaya INH. Machine learning model for umbilical cord classification using combination coiling index and texture feature based on 2-D Doppler ultrasound images. *Health Inf J*. 2022;28(1):14604582221084211. https://doi.org/10.1177/14604582221084211
130. Smail LC, Dhindsa K, Braga LH, Becker S, Sonnadara RR. Using deep learning algorithms to grade hydronephrosis severity: toward a clinical adjunct. *Front Pediatr*. 2020;8:1. https://doi.org/10.3389/fped.2020.00001
131. Weerasinghe NH, Lovell NH, Welsh AW, Stevenson GN. Multiparametric fusion of 3D power Doppler ultrasound for fetal kidney segmentation using fully convolutional neural networks. *IEEE J Biomed Health Inform*. 2021;25(6):2050-2057. https://doi.org/10.1109/JBHI.2020.3027318
132. Yang Y, Yang P, Zhang B. *Automatic Segmentation in Fetal Ultrasound Images Based on Improved U-Net*. IOP Publishing; 2020:012183.
133. Gomez A, Zimmer V, Toussaint N, et al. Image reconstruction in a manifold of image patches: application to whole-fetus ultrasound imaging. In: *Machine Learning for Medical Image Reconstruction*. 2019:226-235. Chapter 21. *Lecture Notes in Computer Science*.
134. Burgos-Artizzu XP, Coronado-Gutierrez D, Valenzuela-Alcaraz B, et al. Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Sci Rep*. 2020;10(1):10200. https://doi.org/10.1038/s41598-020-67076-5
135. Lei B, Zhuo L, Chen S, Li S, Ni D, Wang T. Automatic recognition of fetal standard plane in ultrasound image. 2014:85-88.
136. Bakker MK, Bergman JEH, Krikov S, et al. Prenatal diagnosis and prevalence of critical congenital heart defects: an international retrospective cohort study. *BMJ Open*. 2019;9(7):e028139. https://doi.org/10.1136/bmjopen-2018-028139
137. Holland BJ, Myers JA, Woods CR, Jr. Prenatal diagnosis of critical congenital heart disease reduces risk of death from cardiovascular compromise prior to planned neonatal cardiac surgery: a meta-analysis. *Ultrasound Obstet Gynecol*. 2015;45(6):631-638. https://doi.org/10.1002/uog.14882
138. Mahle WT, Clancy RR, McGaurn SP, Goin JE, Clark BJ. Impact of prenatal diagnosis on survival and early neurologic morbidity in neonates with the hypoplastic left heart syndrome. *Pediatrics*. 2001;107(6):1277-1282. https://doi.org/10.1542/peds.107.6.1277
139. Berhan Y. Predictors of perinatal mortality associated with placenta previa and placental abruption: an experience from a low income country. *J Pregnancy*. 2014;2014:307043. https://doi.org/10.1155/2014/307043
140. Vidaeff AC, Belfort MA, Kemp MW, et al. Updating the balance between benefits and harms of antenatal corticosteroids. *Am J Obstet Gynecol*. 2022;14(2):129-132. https://doi.org/10.1016/j.ajog.2022.10.002
141. Räikkönen K, Gissler M, Kajantie E. Associations between maternal antenatal corticosteroid treatment and mental and behavioral disorders in children. *JAMA*. 2020;323(19):1924-1933. https://doi.org/10.1001/jama.2020.3937

How to cite this article: Horgan R, Nehme L, Abuhamad A. Artificial intelligence in obstetric ultrasound: a scoping review. *Prenat Diagn*. 2023;43(9):1176-1219. https://doi.org/10.1002/pd.6411