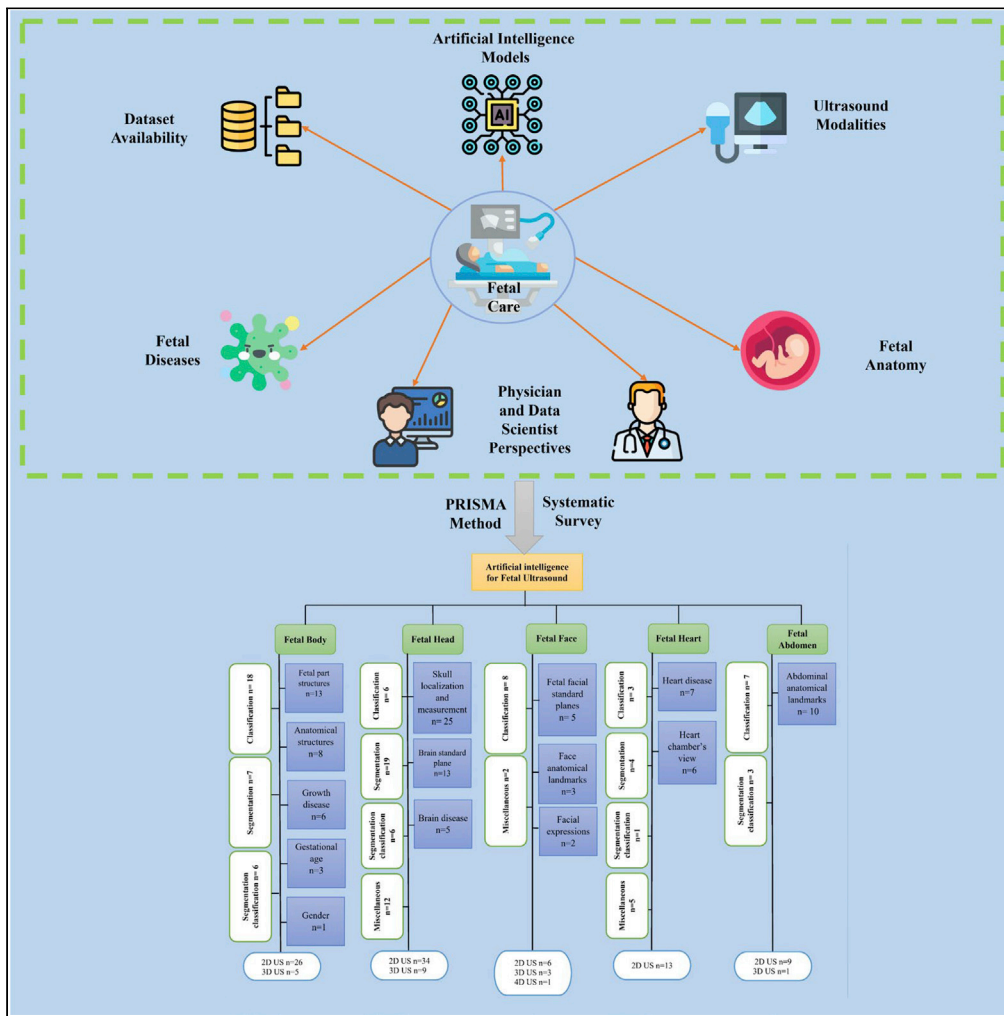


Article

Toward deep observation: A systematic survey on artificial intelligence techniques to monitor fetus via ultrasound images



Mahmood Alzubaidi, Marco Agus, Khalid Alyafei, ..., Mohammed Anbar, Michel Makhlouf, Mowafa Househ

maal28902@hbku.edu.qa (M.A.)
mhhouseh@hbku.edu.qa (M.H.)

Highlights
Artificial intelligence studies to monitor fetal development via ultrasound images

Fetal issues categorized based on four categories — general, head, heart, face, abdomen

The most used AI techniques are classification, segmentation, object detection, and RL

The research and practical implications are included.

Alzubaidi et al., iScience 25, 104713
August 19, 2022 © 2022
<https://doi.org/10.1016/j.isci.2022.104713>



Article

Toward deep observation: A systematic survey on artificial intelligence techniques to monitor fetus via ultrasound images

Mahmood Alzubaidi,^{1,5,*} Marco Agus,¹ Khalid Alyafei,^{2,3} Khaled A. Althelaya,¹ Uzair Shah,¹ Alaa Abd-Alrazaq,^{1,2} Mohammed Anbar,⁴ Michel Makhoulouf,³ and Mowafa Househ^{1,*}

SUMMARY

Several reviews have been conducted regarding artificial intelligence (AI) techniques to improve pregnancy outcomes. But they are not focusing on ultrasound images. This survey aims to explore how AI can assist with fetal growth monitoring via ultrasound image. We reported our findings using the guidelines for PRISMA. We conducted a comprehensive search of eight bibliographic databases. Out of 1269 studies 107 are included. We found that 2D ultrasound images were more popular (88) than 3D and 4D ultrasound images (19). Classification is the most used method (42), followed by segmentation (31), classification integrated with segmentation (16) and other miscellaneous methods such as object-detection, regression, and reinforcement learning (18). The most common areas that gained traction within the pregnancy domain were the fetus head (43), fetus body (31), fetus heart (13), fetus abdomen (10), and the fetus face (10). This survey will promote the development of improved AI models for fetal clinical applications.

INTRODUCTION

Background

Artificial intelligence (AI) is a broad discipline that aims to replicate the inherent intelligence shown by people via artificial methods (Hassabis et al., 2017). Recently, AI techniques have been widely utilized in the medical sector (Miller and Brown, 2018). Historically, AI techniques were standalone systems with no direct link to medical imaging. With the development of new technology, the idea of 'joint decision-making' between people and AI offers the potential of boosting high performance in the area of medical imaging (Savadjiev et al., 2019).

In computer science, machine learning (ML), deep learning (DL), artificial neural network (ANN) and reinforcement learning (RL) are subset techniques of AI that are used to perform different tasks on medical images such as classification, segmentation, object identification, and regression (Fatima and Pasha, 2017; Kim et al., 2019b; Shahid et al., 2019). Diagnosis using computer-aided detection (CAD) has moved toward becoming AI automated process in the medical images (Castiglioni et al., 2021), which include most of the medical imaging data such as X-ray radiography, fluoroscopy, MRI, medical ultrasonography or ultrasound, endoscopy, elastography, tactile imaging, and thermography (Alzubaidi et al., 2021a, 2021b; Fujita, 2020). However, digitized medical images come with a plethora of new information, possibilities, and challenges. Therefore, AI techniques are able to address some of these challenges by showing impressive accuracy and sensitivity in identifying imaging abnormalities. These techniques promise to enhance tissue-based detection and characterization with the potential to improve diagnoses of diseases (Tang, 2020).

At present, the use of AI techniques in medical images has been discussed in depth across many medical disciplines, including identifying cardiovascular abnormalities, detecting fractures and other musculoskeletal injuries, aiding in diagnosing neurological diseases, reducing thoracic complications and conditions, screening for common cancers, and many other prognoses and diagnosis tasks (Castiglioni et al., 2021; Cheikh et al., 2020; Deo, 2015; Handelman et al., 2018; Miotto et al., 2017). Furthermore, AI techniques have shown the ability to provide promising findings when utilizing prenatal medical images, such as

¹College of Science and Engineering, Hamad Bin Khalifa University, Member of Qatar Foundation, Doha, Qatar

²Weil Cornell Medical College-Qatar, Doha, Qatar

³Sidra Medical and Research Center, Member of Qatar Foundation, Doha, Qatar

⁴National Advanced IPv6 Centre, University Sains Malaysia, Penang, Malaysia

⁵Lead contact

*Correspondence: maal28902@hbku.edu.qa (M.A.), mhouseh@hbku.edu.qa (M.H.)

<https://doi.org/10.1016/j.isci.2022.104713>



monitoring fetal development at each stage of pregnancy, predicting a pregnant placenta's health, and identifying potential complications (Balayla and Shrem, 2019; Chen et al., 2021a; Iftikhar et al., 2020; Raef and Ferdousi, 2019). AI techniques may assist with detecting several fetal diseases and adverse pregnancy outcomes with complex etiologies and pathogeneses such as amniotic band syndrome, congenital diaphragmatic hernia, congenital high airway obstruction syndrome, fetal bowel obstruction, gastroschisis, omphalocele, pulmonary sequestration, and sacrococcygeal teratoma (Correa et al., 2016; Larsen et al., 2013; Sen, 2017). Therefore, more research is needed to understand the role of utilizing AI techniques in the early stage of pregnancy, both to prevent and reduce unfavorable outcomes as well as provide an understanding of fetal anomalies and illnesses throughout pregnancy and for drastically reducing the need of more invasive diagnostic procedures that may be harmful for the fetus (Chen et al., 2021b).

Various imaging modalities (e.g., ultrasound, MRI, and computed tomography (CT)) are available and can be utilized by AI techniques during pregnancy (Kim et al., 2019b). In medical literature, ultrasound imaging has become popular and is used during all pregnancy trimesters. Ultrasound is crucial for making diagnoses, along with tracking fetal growth and development. In addition, ultrasound can provide both precise fetal anatomical information as well as high-quality photos and increased diagnostic accuracy (Avola et al., 2021; Chen et al., 2021a). There are numerous benefits and few limitations when using ultrasound. Acquiring the device is a very inexpensive process, particularly when compared to other instruments used on the same organs, such as MRI or positron emission tomography (PET). Unlike other acquisition equipment for comparable purposes, ultrasound scanners are portable and long-lasting. The second major benefit is that unlike MRI and CT the ultrasound machine does not pose any health concerns for pregnant women as the produced signals are entirely safe for both the mother and the fetus (Whitworth et al., 2015). Although the literature exploring the benefits of utilizing AI technologies for ultrasound imaging diagnostics in prenatal care has grown, more research is needed to understand the different roles AI can have in diagnostic prenatal care using ultrasound images.

Research objective

Previous reviews on ultrasound medical images for pregnant women using AI techniques have not been thoroughly conducted. Early reviews have attempted to summarize the use of AI in the prenatal period (Al-yousif et al., 2021; Arjunan and Thomas, 2020; Avola et al., 2021; Davidson and Boland, 2021; Garcia-Canadilla et al., 2020; Hartkopf et al., 2019; Kaur et al., 2017; Komatsu et al., 2021a; Shiney et al., 2017; Torrents-Barrena et al., 2019). However, these reviews are not comprehensive because they only target specific issues such as fetal cardiac function (Garcia-Canadilla et al., 2020), Down syndrome (Arjunan and Thomas, 2020), and fetal CNS (Hartkopf et al., 2019). Similarly, other reviews have not focused on AI techniques and ultrasound images as the primary intervention (Kaur et al., 2017; Shiney et al., 2017; Torrents-Barrena et al., 2019). Two reviews covered the use of general AI techniques utilized in the prenatal period through ultrasound. However, they did not focus on the fetus as the primary population in their review (Avola et al., 2021; Davidson and Boland, 2021). Another systematic review (Al-yousif et al., 2021) addressed AI classification technologies for fetal health, but it only targeted cardiotocography disease. Other review (Komatsu et al., 2021a) they introduced various areas of medical AI research including obstetrics but the main focus on US imaging in general. Therefore, a comprehensive survey is necessary to understand the role of AI technologies for the entire prenatal care period using ultrasound images. This work will be the first comprehensive study that systematically reviews how AI techniques can assist during pregnancy screenings of fetal development and improve fetus growth monitoring. In this survey we aim to: (1) Describe the application of AI techniques on different fetal spatial ultrasound dimensions; (2) Discuss how different methodologies are used—classification, segmentation, object-detection, regression, and RL, (3) Highlight the dataset acquisition and availability; (4) Identify practical and research implications and gaps for future research.

RESULTS

Search results

As shown in Figure 1, the search yielded a total of 1269 citations. After excluding 457 duplicates, there were 812 unique titles and abstracts. A total of 598 citations were excluded after evaluating the titles and abstracts. After full-text screening, $n = 107$ citations were excluded from the remaining $n = 214$ papers. The narrative synthesis includes a total of 107 studies. Across screening steps, both authors reported the excluding reason for each study to ensure reliability of the work. The total excluded studies were $n = 704$ articles. These studies did not meet selection criteria because of the following reasons: (1) irrelevant,

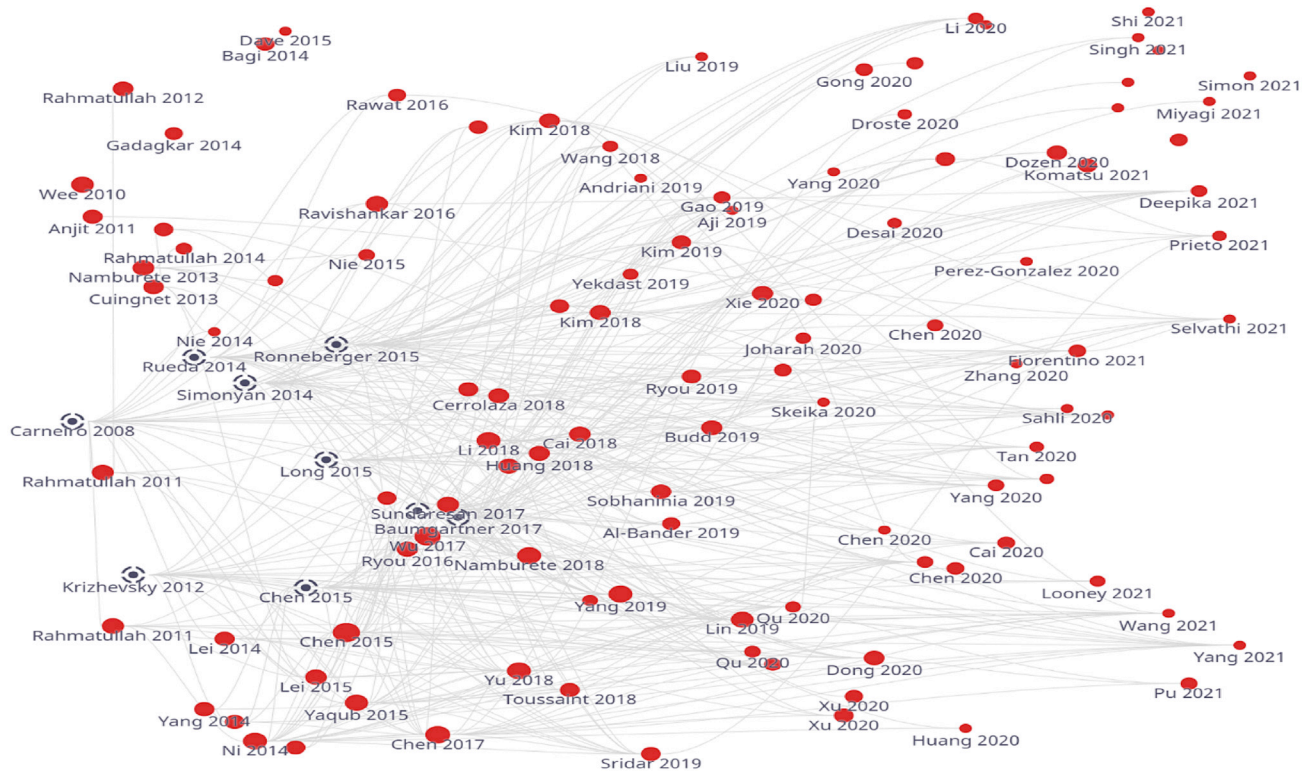


Figure 1. Literature map showing the selected studies in red dot and recommended studies in black dot

(2) wrong intervention, (3) wrong population, (4) wrong publication type, (5) unavailable, and (6) foreign language article. These terminologies are defined in Table 1.

Assessment of bibliometrics analysis

To validate the selected studies, we used a bibliometrics analysis (Kokol et al., 2021) tool that explores the literature through citations and automatically recommends highly related articles to our scope. This ensured that we were unlikely to miss any relevant study. Figure 1 provides a citations map between selected studies over time. The map shows that all selected studies are relevant over time from 2010 to 2021 and recommends nine studies that may be relevant to our scope based on the interactive citation. However, these studies were excluded because they did not meet our inclusion criteria. Therefore, we concluded that our search was comprehensive and included most of the relevant studies.

Identification of result themes

In Figure 2, we grouped the studies by the fetus organ that each study addresses including: 1) Fetus head (n = 43, 40.18%), 2) fetus body (n = 31, 28.97%), 3) fetus heart (n = 13, 12.14%), 4) fetus face (n = 10, 9.34%), and 5) fetus abdomen (n = 10, 9.34%). Each group is classified by sub-group as detailed in Figure 3. In addition, most of the studies used 2D US (n = 88, 82.24%) and 3D/4D US were rare and reported only in (n = 19, 17.75%). All tasks are grouped based on the most common use. We found that classification was used in (n = 42, 39.25%), segmentation was used in (n = 31, 28.97%), both classification and segmentation were used together in (n = 16, 14.95%), object-detection, regression, and RL were seen as miscellaneous in (n = 18, 16.82%).

Definition of result themes

Ultrasound imaging modalities

There was a total of (n = 88, 82.24%) studies that utilized 2D US images. Therefore, here is a brief about what the 2D ultrasound is. Sound waves are used in all ultrasounds to produce an image. The conventional ultrasound picture of a growing fetus is a 2D image. The fetus's internal organs may be seen with a 2D ultrasound, which provides outlines and flat-looking pictures. 2D ultrasounds have been widely available for a

Table 1. Definition of the excluded terminologies

| Terminologies | Definition |
|------------------------|--|
| Irrelevant | <ul style="list-style-type: none"> Publication that is not related to the scope of this survey |
| Wrong intervention | <ul style="list-style-type: none"> Publication that targets fetus health but not using AI technology and ultrasound image |
| Wrong population | <ul style="list-style-type: none"> Publication that uses AI technology and ultrasound image but did not target fetus health |
| Wrong publication type | <ul style="list-style-type: none"> Publication that is conference abstract, review, magazine, or newspaper |
| Unavailable | <ul style="list-style-type: none"> Publication that is not accessible or cannot be found |
| Foreign language | <ul style="list-style-type: none"> Publication that is not written in English |

long time and have a very good safety record. These devices do not pose the same dangers as X-rays because they employ non-ionizing rather than ionizing radiation (Avola et al., 2021). Ultrasounds are typically conducted at least once throughout pregnancy, most often between 18 and 22 weeks in the second trimester (Bethune et al., 2013). This examination, also known as a level two ultrasound or anatomy scan, is performed to monitor the fetus's development. Ultrasounds may be used to examine a variety of things during pregnancy, including: 1) How the fetus is developing; 2) The gestational age of the fetus; 3) Any abnormalities within the uterus, ovaries, cervix, or placenta; 4) The number of fetuses you are carrying; 5) Any difficulties the fetus may be experiencing; 6) The fetus's heart rate; 7) Fetal growth and location in the uterus, 8) Amniotic fluid level, 9) Sex assignment, 10) Signs of congenital defects, and 11) Signs of Down syndrome (Driscoll and Gross, 2009; Masselli et al., 2014).

There was a total of (n = 18, 16.8%) studies that utilized 3D. Therefore, here is a brief about what is 3D ultrasound. Observing three-dimensional (3D) ultrasound images has increasingly grown in recent years. Although 3D ultrasounds may be beneficial in identifying facial or skeletal abnormalities (Huang and Zeng, 2017), 2D ultrasounds are commonly utilized in medical settings because they can clearly reveal the interior organs of a growing fetus. The picture of a 3D ultrasound is created by putting together several 2D photos obtained from various angles. Many parents like 3D photos because they believe they can see their baby's face more clearly than they can with flat 2D images.

Despite this, the Food and Drug Administration (FDA) does not recommend using a 3D ultrasound for entertainment purposes (Abramowicz, 2010). The "as low as reasonably attainable" (ALARA) approach directs ultrasound technologists to be used solely in a clinical setting in order to reduce exposure to heat and radiation (Abramowicz, 2015). Although an ultrasound is generally thought to be harmless, there is insufficient data to determine what long-term exposure to ultrasound may be to a fetus or a pregnant woman. There is no way of knowing how long a session will take or if the ultrasound equipment will work correctly in non-clinical situations such as places that give "keepsake" photos (Dowdy, 2016).

4D ultrasound is identical to a 3D ultrasound, but the image it creates is updated constantly – comparable to a moving image. This sort of ultrasound is usually performed for fun rather than for medical concerns. As previously mentioned, because ultrasound is medical equipment that should only be used for medical purposes (Kurjak et al., 2007), the FDA does not advocate getting ultrasounds for enjoyment or bonding purposes. Unless their doctor or midwife has recommended it as part of prenatal care, pregnant patients should avoid non-medical locations that provide ultrasounds (Dowdy, 2016). Although there is no proven risk, both 3D and 4D ultrasounds employ higher-than-normal amounts of ultrasound energy and the sessions are longer compared to 2D which may have adverse fetal effects (Mack, 2017; Payan et al., 2018).

AI image processing task

There was a total of (n = 42, 39.25%) studies that utilized classification, which is the process of assigning one or more labels to an image, and it is one of the most fundamental problems concerning accurate computer

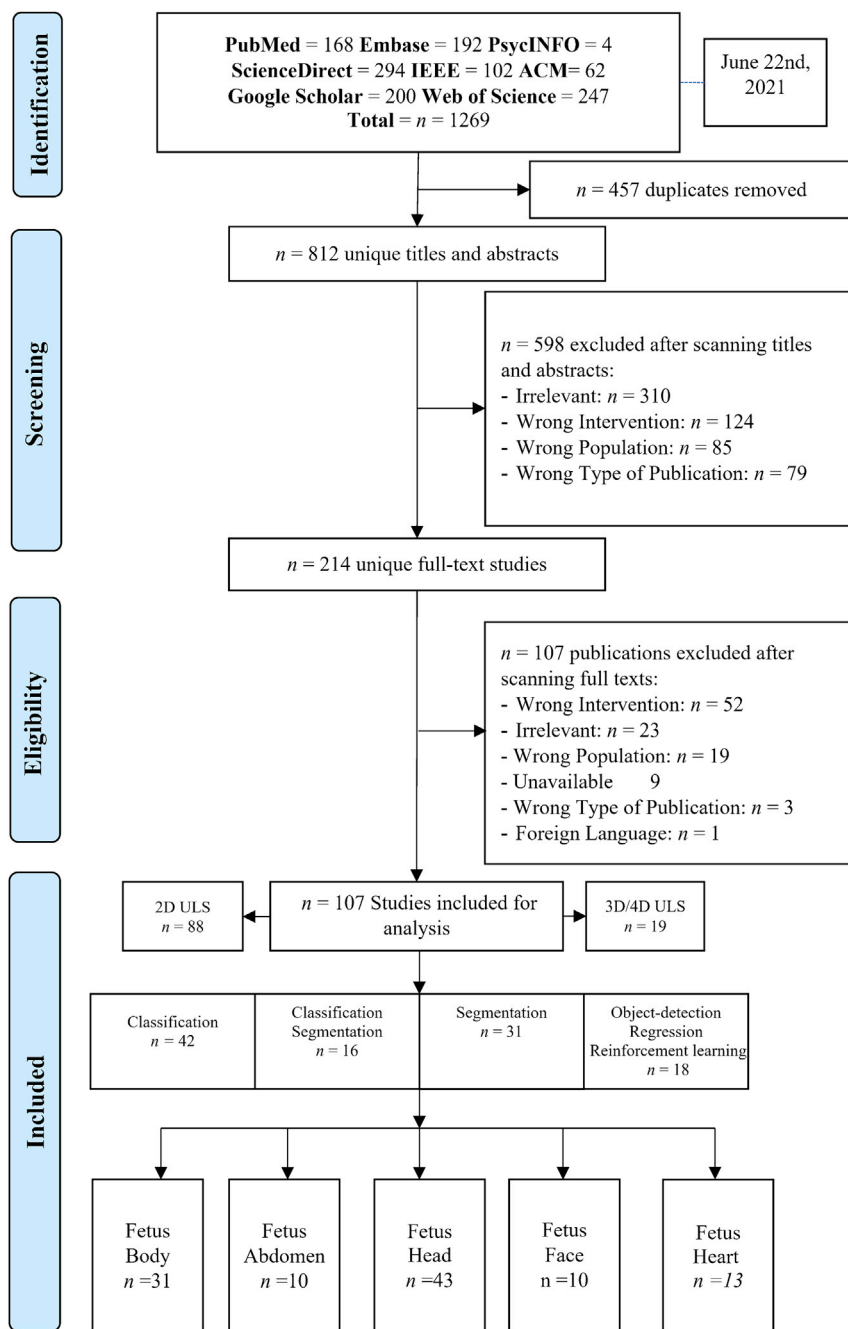


Figure 2. PRISMA diagram showing our literature search inclusion process

vision and pattern recognition. It has many applications including image and video retrieval, video surveillance, web content analysis, human-computer interaction, and biometrics. Feature coding is an important part of image classification, and several coding methods have been developed in recent years. In general, image classification entails extracting picture characteristics before classifying them (Naeem and Bin-Salem, 2021). As a result, the main aspect of utilizing image classification is understanding how to extract and evaluate image characteristics (Wang et al., 2020).

There was a total of ($n = 31$, 28.97%) studies that utilized segmentation, which is one of the most complex tasks in medical image processing involving distinguishing the pixels of organs or lesions from background

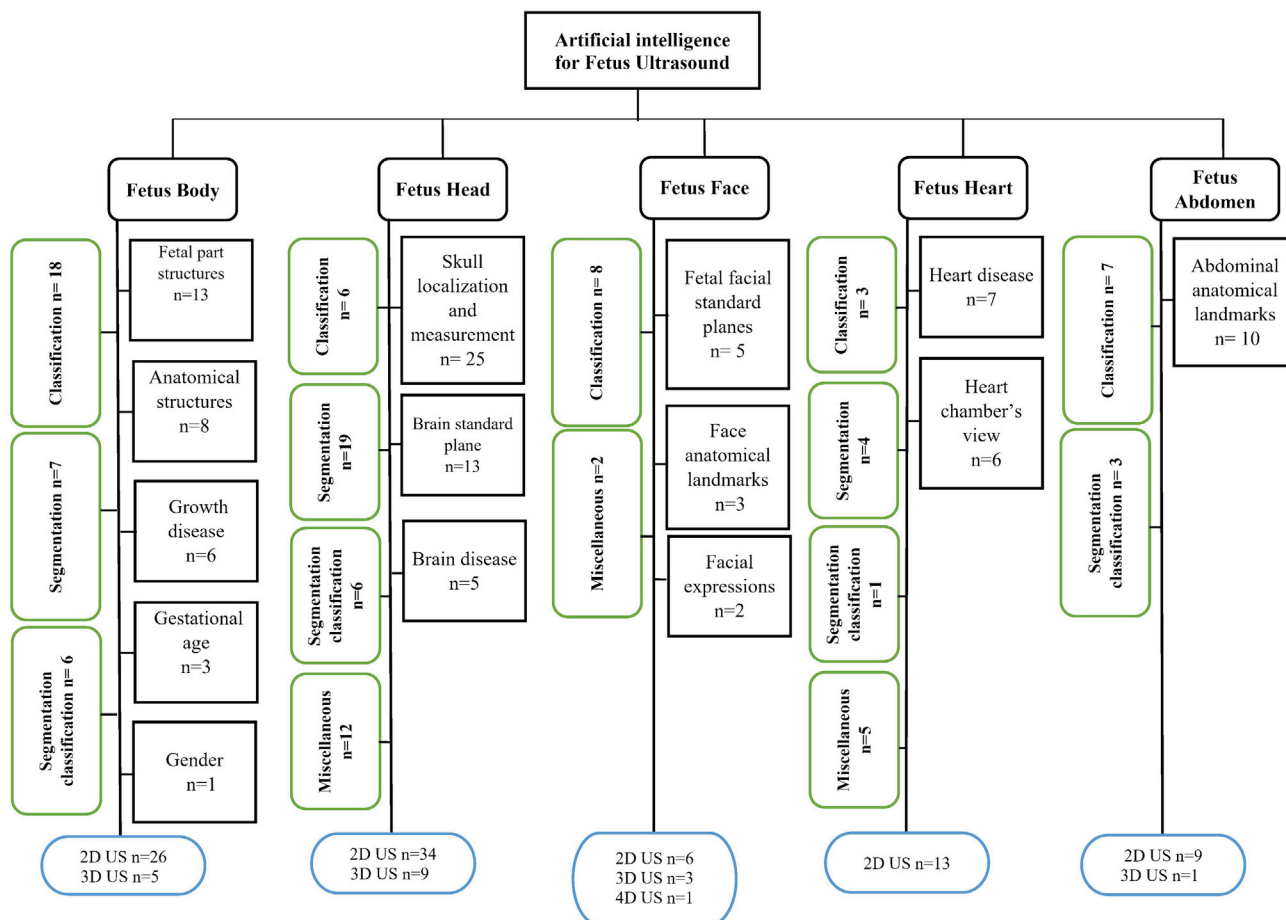


Figure 3. Summary of AI methods implemented on fetal ultrasound images

medical images such as CT or MRI scans (Bin-Salem et al., 2022). A number of researchers have suggested different automatic segmentation methods by using existing technology (Asgari Taghanaki et al., 2021). Previously, traditional techniques such as edge detection filters and mathematical algorithms were used to construct earlier systems (Bali and Singh, 2015). After this, machine learning techniques for extracting hand-crafted characteristics were the dominating method for a substantial time period. The main issue for creating such a system has always been designing and extracting these characteristics, and the complexity of these methods has been seen as a substantial barrier to deployment. In the 2000s, deep learning methods began to grow in popularity because of advancements in technology because this method had significantly better skills in image processing jobs. Deep learning methods have emerged as a top choice for image segmentation, particularly medical image segmentation, because of their promising capabilities (Hesamian et al., 2019).

There was a total of (n = 12, 11.21%) studies that utilized object detection techniques, which is the process of locating and classifying items. A detection method is used in biomedical images to determine the regions where the patient's lesions are situated as box coordinates. There are two kinds of deep learning-based object detection. These region proposal-based algorithms are one example. Using a selective search technique, this method extracts different kinds of patches from input images. After that, the trained model determines if each patch contains numerous items and classifies them according to their region of interest (ROI). In particular, the region proposal network was created to speed up the detection process. Object identification in the other methods is done using the regression-based algorithm as a one-stage network. These methods use image pixels to directly identify and detect bounding box coordinates and class probabilities within entire images (Kim et al., 2019b).

Table 2. Characteristics of the included studies (n = 107)

| Characteristics | Number of studies |
|--|-------------------|
| Type of publication | (n, %) |
| Journal article | (53, 49.53) |
| Conference proceedings | (43, 40.18) |
| Book chapter | (11, 10.28) |
| Year of publication | |
| 2021 | (19, 17.75) |
| 2020 | (30, 28.03) |
| 2019 | (14, 13.08) |
| 2018 | (11, 10.28) |
| 2017 | (6, 5.60) |
| 2016 | (4, 3.73) |
| 2015 | (6, 5.60) |
| 2014 | (9, 8.41) |
| 2013 | (3, 2.80) |
| 2012 | (1, 0.93) |
| 2011 | (3, 2.80) |
| 2010 | (1, 0.93) |
| Fetal Organ | |
| Fetus body | (31, 28.97) |
| Fetal part structures | (13, 12.14) |
| Anatomical structures | (8, 7.47) |
| Growth disease | (6, 5.60) |
| Gestational age | (3, 2.80) |
| Gender | (1, 0.93) |
| Fetus head | (43, 40.18) |
| Skull localization and measurement | (25, 23.36) |
| Brain standard plane | (13, 12.14) |
| Brain disease | (5, 4.67) |
| Fetus face | (10, 9.34) |
| Fetal facial standard planes | (5, 4.67) |
| Face anatomical landmarks | (3, 2.80) |
| Facial expressions | (2, 1.86) |
| Fetus heart | (13, 12.14) |
| Heart disease | (7, 6.54) |
| Heart chambers view | (6, 5.60) |
| Fetus abdomen | (10, 9.34) |
| Abdominal anatomical landmarks | (10, 9.34) |
| Publication Country and top institute | |
| China | (41, 38.31) |
| Shenzhen University | (13, 12.14) |
| Beihang University | (5, 4.67) |
| Chinese University of Hong Kong | (4, 3.73) |
| Hebei University of Technology | (2, 1.86) |
| Fudan University | (2, 1.86) |
| Shanghai Jiao Tong University | (2, 1.86) |

(Continued on next page)

Table 2. Continued

| Characteristics | Number of studies |
|--------------------------------------|-------------------|
| South China University of Technology | (2, 1.86) |
| Other institutes | (11, 10.28) |
| UK | (25, 23.36) |
| University of Oxford | (20, 18.69) |
| Imperial College London | (4, 3.73) |
| King's College, London | (1, 0.93) |
| India | (14, 13.08) |
| Japan | (4, 3.73) |
| Indonesia | (4, 3.73) |
| USA | (3, 2.80) |
| South Korea | (3, 2.80) |
| Iran | (3, 2.80) |
| Australia | (1, 0.93) |
| Canada | (1, 0.93) |
| Mexico | (1, 0.93) |
| France | (1, 0.93) |
| Italy | (1, 0.93) |
| Tunisia | (1, 0.93) |
| Spain | (1, 0.93) |
| Iraq | (1, 0.93) |
| Brazil | (1, 0.93) |
| Malaysia | (1, 0.93) |

There was a total of (n = 2, 1.8%) studies that utilized RL. The concept behind RL is that an artificial agent learns by interacting with its surroundings. It enables agents to autonomously identify the best behavior to exhibit in each situation to optimize performance on specified metrics. The basic concept underlying RLs made up of many components. The RL agent is the process's decision-maker, and it tries to perform an action that has been recognized by the environment. Depending on the activity performed, the agent receives a reward or penalty from its surroundings. Exploration and exploitation are used by RL agents to determine which activities provide the most rewards. In addition, the agent obtains information about the condition of the environment (Sahba et al., 2006). For medical image analysis applications, RL provides a powerful framework. RL has been effectively utilized in a variety of applications, including landmark localization, object detection, and registration using image-based parameter inference. RL has also been shown to be a viable option for handling complex optimization problems such as parameter tuning, augmentation strategy selection, and neural architecture search. Existing RL applications for medical imaging can be split into three categories: parametric medical image analysis, medical image optimization, and numerous miscellaneous applications (Zhou et al., 2021).

Description of included studies

As shown in (Table 2), more than half of the studies (n= 53, 49.53%) were obtained from journal articles whereas the other half were found in conference proceedings (n= 43, 40.18%) and book chapters (n = 11, 10.28%). Studies were published between 2010 and 2021. The majority of studies were published in 2020 (n = 30, 28.03%), followed by 2021 (n = 19, 17.75%), 2019 (n = 14, 13.08%), 2018 (n = 11, 10.28%), 2014 (n = 9, 8.41%), 2017 (n = 6, 5.60%), 2015 (n = 6, 5.60%), 2016 (n = 4, 3.73%), 2013 (n = 3, 2.80%), 2011 (n = 3, 2.80%), 2012 (n = 1, 0.93%), and 2010 (n = 1, 0.93%). In regard to studies examining fetal organs, fetus skull localization and measurement was reported on in (n = 25, 23.36%) studies, followed by fetal part structures (n = 13, 12.14%) and brain standard plane in (n = 13, 12.14%) studies. Abdominal anatomical landmarks were reported in (n = 10, 9.34%) studies. In addition, fetus body anatomical structures were reported in (n = 8, 7.47%) studies, heart disease was reported on in (n = 7, 6.54%) studies, and growth disease (n = 6, 5.60%) and brain disease were also reported in (n = 5, 4.67%) studies. The view of heart chambers was

reported on in ($n = 6$, 5.60%) studies, fetal facial standard planes were reported in ($n = 5$, 4.67%) studies, gestational age ($n = 3$, 2.80%) and face anatomical landmarks were reported on in ($n = 3$, 2.80%) studies, facial expressions were reported on in ($n = 2$, 1.86%) studies, and gender identification was reported on only in ($n = 1$, 0.93%) study. In addition, most of the studies were conducted in China ($n = 41$, 38.31%), followed by the UK ($n = 25$, 23.36%), and India ($n = 14$, 13.08%). Surprisingly, some institutes mainly focused on this field and contributed a high number of studies. For example, many ($n = 20$, 18.69%) studies were conducted at the University of Oxford, UK and several ($n = 13$, 12.14%) were conducted in Shenzhen University, China.

Artificial intelligence for fetus ultrasound

Fetus body

As shown in (Table 2), the overall purpose of several ($n = 31$, 28.97%) studies was to identify general characters of the fetus (e.g., gestational age, gender) or diseases such as intrauterine growth restriction (IUGR). In addition, most of the studies aimed to identify the fetus itself in the uterus or fetus part structure during different trimesters. The first trimester is from week 1 to the end of week 12. The second trimester is from week 13 to the end of week 26, and the third trimester is from week 27 to the end of the pregnancy. We also report on a performance comparison between the various techniques for each fetus body group including objective, backbone methods, optimization, fetal age, best obtained result, and observations (see Table 3).

Fetal part structures

Classification. For identification and localization of fetal part structure, image classification was widely used in ($n = 9$, 8.41%) studies. Only 2D US images were utilized for classification purposes. Studies (Gao and Noble, 2019; Maraci et al., 2015; Yaqub et al., 2015) used a classification task to identify and locate fetal skull, heart, and abdomen from 2D US images in the second trimester. Besides that, in (Cai et al., 2018), classification was used to locate the exact abdominal circumference plane (ACP), and in (Cai et al., 2020), more results were obtained by locating head circumference plane (HCP), the abdominal circumference plane (ACP), and the femur length plane (FLP). Furthermore, in studies (Chen et al., 2015, 2017; Sridar et al., 2019) researching the second and third trimester, multi-classes were used to identify and locate various parts including (1) fetal abdominal standard plane (FASP); (2) fetal face axial standard plane (FFASP); and (3) fetal four-chamber view standard plane (FFVSP) of the heart. Study (Burgos-Artiztu et al., 2020) located the mother's cervix in addition to abdomen, brain, femur, and thorax.

Segmentation. Image segmentation task was used in ($n = 3$, 2.80%) studies for the purpose of fetal part structure identification. 2D US imaging (Liu et al., 2021) was used to segment neonatal hip bone including seven key structures: (1) chondro-osseous border (CB), (2) femoral head (FH), (3) synovial fold (SF), (4) joint capsule and perichondrium (JCP), (5) labrum (La), (6) cartilaginous roof (CR), and (7) bony roof (BR). Only one study in (Weerasinghe et al., 2021) provided segmentation model that was able to segment the fetus kidney using 3D US in the third trimester. Segmentation was used to locate fetal head, femur, and humerus using 2D US as seen in (Rahmatullah et al., 2014).

Classification and segmentation. In (Ryou et al., 2019), whole fetal segmentation followed by classification tasks were used to locate the head, abdomen (in sagittal view), and limbs (in axial view) using 3D US images taken in the first trimester.

Anatomical part structures

Classification. For identification and localization of anatomical structures, image classification was used in ($n = 4$, 3.73%) studies. Only one study (Ryou et al., 2016) used 3D US images and other studies used 2D US (Ravishankar et al., 2016; Toussaint et al., 2018; Wee et al., 2010). In (Ryou et al., 2016), the whole fetus was localized in the sagittal plane and classifier was then applied to the axial images to localize one of three classes (head, body and non-fetal) during the second trimester. Multi-classification methods were utilized in (Toussaint et al., 2018) to localize head, spine, thorax, abdomen, limbs, and placenta in the third trimester. Moreover, binary classification was used in (Ravishankar et al., 2016) to identify abdomen versus non-abdomen. In (Wee et al., 2010), a classification method was used to detection and measure nuchal translucency (NT) in the beginning of second trimester. Moreover, monitoring NT combined with maternal age can provide effective insight of screening Down syndrome.

Table 3. Articles published using AI to improve fetus body monitoring: objective, backbone methods, optimization, fetal age, and AI tasks

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|-------------------------------|--|---|--|-----------------------|-----------------------------|
| Fetal Part Structures | | | | | |
| (Maraci et al., 2015) | To identify the fetal skull, heart and abdomen from ultrasound images | SVM as the classifier | Gaussian Mixture Model (GMM) Fisher Vector (FV) | 26 th week | classification |
| (Liu et al., 2021) | To segment the seven key structures of the neonatal hip joint | Neonatal Hip Bone Segmentation Network (NHBSNet) | Feature Extraction Module Enhanced Dual Attention Module (EDAM) Two-Class Feature Fusion Module (2-Class FFM) Coordinate Convolution Output Head (CCOH) | 16 - 25 weeks. | segmentation |
| (Rahmatullah et al., 2014) | To segment organs head, femur, and humerus in ultrasound images using multilayer super pixel images features | Simple Linear Iterative Clustering (SLIC) Random forest | Unary pixel shape feature image moment | N/A | segmentation |
| (Weerasinghe et al., 2021) | To automate kidney segmentation using fully convolutional neural networks. | FCNN: U-Net & UNET++ | N/A | 20 to 40 weeks | segmentation |
| (Burgos-Artizzu et al., 2020) | To evaluate the maturity of current Deep Learning classification techniques for their application in a real maternal-fetal clinical environment | CNN DenseNet-169 | N/A | 18 to 40 weeks | classification |
| (Cai et al., 2020) | To use the learnt visual attention maps to guide standard plane detection on all three standard biometry planes: ACP, HCP and FLP. | Temporal SonoEyeNet (TSEN) Temporal attention module: Convolutional LSTM Video classification module: Recurrent Neural Networks (RNNs)+ | CNN feature extractor: VGG-16 | N/A | classification |
| (Ryou et al., 2019) | To support first trimester fetal assessment of multiple fetal anatomies including both visualization and the measurements from a single 3D ultrasound scan | Multi-Task Fully Convolutional Network (FCN) U-Net | N/A | 11 to 14 weeks | Segmentation Classification |

(Continued on next page)

Table 3. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|------------------------------|---|---|--|------------------------|----------------|
| (Sridar et al., 2019) | To automatically classify 14 different fetal structures in 2D fetal ultrasound images by fusing information from both cropped regions of fetal structures and the whole image | support vector machine (SVM)+ Decision fusion | Fine-tuning AlexNet CNN | 18 to 20 weeks | Classification |
| (Chen et al., 2017) | To automatic identification of different standard planes from US images | T-RNN framework: LSTM | Features extracted using J-CNN classifier | 18 to 40 weeks | Classification |
| (Cai et al., 2018) | To classify abdominal fetal ultrasound video frames into standard AC planes or background | M-SEN architecture Discriminator CNN | Generator CNN | N/A | Classification |
| (Gao and Noble, 2019) | To detect multiple fetal structures in free-hand ultrasound | CNN Attention Gated LSTM | Class Activation Mapping (CAM) | 28 to 40 weeks | classification |
| (Yaqub et al., 2015) | To extract features from regions inside the images where meaningful structures exist. | Guided Random Forests | Probabilistic Boosting Tree (PBT) | 18 to 22 weeks | Classification |
| (Chen et al., 2015) | To detect standard planes from US videos | T-RNN LSTM (Transferred RNN) | Spatio-Semporal Feature J-CNN | 18 to 40 weeks | Classification |
| Anatomical Structures | | | | | |
| (Yang et al., 2019) | To propose the first and fully automatic framework in the field to simultaneously segment fetus, gestational sac and placenta, | 3D FCN + RNN hierarchical deep supervision mechanism (HiDS) | BiLSTM module denoted as FB-nHiDS | 10 - 14 weeks | Segmentation |
| (Looney et al., 2021) | To segment the placenta, amniotic fluid, and fetus. | FCNN | N/A | 11 - 19 weeks | Segmentation |
| (Li et al., 2017) | To segment the amniotic fluid and fetal tissues in fetal US images | The encoder-decoder network based on VGG16 | N/A | 22 ND weeks | Segmentation |

(Continued on next page)

Table 3. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|-----------------------------------|---|---|--|----------------|-----------------------------|
| (Ryou et al., 2016) | To localize the fetus and extract the best fetal biometry planes for the head and abdomen from first trimester 3D fetal US images | CNN | Structured Random Forests | 11 - 13 weeks | Classification |
| (Toussaint et al., 2018) | To detect and localize fetal anatomical regions in 2D US images | ResNet18 | Soft Proposal Layer (SP) | 22 - 32 weeks | Classification |
| (Ravishankar et al., 2016) | To reliably estimate abdominal circumference | CNN + Gradient Boosting Machine (GBM) | Histogram of Oriented Gradient (HoG) | 15 - 40 weeks | Classification |
| (Wee et al., 2010) | To detect and recognize the fetal NT based on 2D ultrasound images by using artificial neural network techniques. | Artificial Neural Network (ANN) | Multilayer Perceptron (MLP) Network Bidirectional Iterations Forward Propagations Method (BIFP) | N/A | Classification |
| (Liu et al., 2019) | To detect NT region | U-Net NT Segmentation PCA NT Thickness Measurement | VGG16 NT Region Detection | 4 - 12 weeks | Segmentation |
| Growth disease | | | | | |
| (Bagi and Shreedhara, 2014) | To propose the biometric measurement and classification of IUGR, using OpenGL concepts for extracting the feature values and ANN model is designed for diagnosis and classification | ANN Radial Basis Function (RBF) | OpenGL | 12–40 Weeks | Classification |
| (Selvathi and Chandralekha, 2021) | To find the region of interest (ROI) of the fetal biometric and organ region in the US image | DCNN AlexNet | N/A | 16 -27 weeks | Classification |
| (Rawat et al., 2016) | To detect fetal abnormality in 2D US images | ANN + Multilayered perceptron neural networks (MLPNN) | Gradient vector flow (GVF) Median Filtering | 14 - 40 weeks | Classification segmentation |
| (Gadagkar and Shreedhara, 2014) | To develop a computer-aided diagnosis and classification tool for extracting ultrasound sonographic features and classify IUGR fetuses | ANN | Two-Step Splitting Method (TSSM) for Reaction-Diffusion (RD) | 12–40 Weeks | Classification segmentation |

(Continued on next page)

Table 3. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|--------------------------------|--|--------------------------------|---|------------------------|-----------------------------|
| (Andriani and Mardhiyah, 2019) | To develop an automatic classification algorithm on the US examination result using Convolutional Neural Network in Blighted Ovum detection | CNN | N/A | N/A | Classification |
| (Yekdast, 2019) | To propose an intelligent system based on combination of ConvNet and PSO for Down syndrome diagnosis. | CNN | Particle Swarm Optimization (PSO) | N/A | Classification |
| (Maraci et al., 2020) | To automatically detect and measure the trans cerebellar diameter (TCD) in the fetal brain, which enables the estimation of fetal gestational age (GA) | CNN FCN | N/A | 16- - 26 weeks | Classification segmentation |
| (Chen et al., 2020a) | To accurately estimate the gestational age from the fetal lung region of US images. | U-NET | N/A | 24 - 40 weeks | Classification segmentation |
| (Prieto et al., 2021) | To classify, segment, and measure several fetal structures for the purpose of GA estimation | U-NET RESTNET | Residual UNET (RUNET) | 16 th weeks | Classification segmentation |
| (Maysanjaya et al., 2014) | To measure the accuracy of Learning Vector Quantization (LVQ) to classify the gender of the fetus in the US image" | ANN | Learning Vector Quantization (LVQ) Moment invariants | N/A | Classification |

Segmentation. An image segmentation task was used in ($n = 4$, 3.73%) studies for the purpose of fetal image segmentation was target task anatomical structures identification. In (Looney et al., 2021; Yang et al., 2019), 3D US were utilized to segment fetus, gestational sac amniotic fluid, and placenta in the beginning of the second trimester. However, in (Li et al., 2017), 2D US was used to segment amniotic fluid and the fetus in the late trimester. Finally, 2D US in (Liu et al., 2019) was utilized to segment and measure NT in the first trimester of pregnancy.

Growth disease diagnosis

Classification. For diagnosis of fetal growth, disease image classification was conducted in ($n = 4$, 3.73%) studies using 2D US. In (Bagi and Shreedhara, 2014) binary classification was used for early diagnosis of intrauterine growth restriction (IUGR) at the third trimester of pregnancy. The features considered to determine a diagnosis of IUGR are gestational age (GA), biparietal diameter (BPD), abdominal circumference (AC), head circumference (HC), and femur length (FL). In addition, binary classification was used in the second trimester of pregnancy to identify normal versus abnormal fetus growth as seen in (Selvathi and Chandralekha, 2021). In (Andriani and Mardhiyah, 2019), binary classification was used to identify the normal growth of the ovum by distinguishing blighted ovum from healthy ovum. Lastly, the binary classification was used to distinguish between healthy fetuses and those with Down syndrome (Yekdast, 2019).

Classification and segmentation. For diagnosis of fetal growth disease, image classification along with segmentation were used in ($n = 2$, 1.86%) studies using 2D US. In (Gadagkar and Shreedhara, 2014; Rawat et al., 2016), segmentation of the region of interest (ROI), followed by classification to diagnosis (IUGR) (normal versus abnormal) were carried out. This was done by using US images taken in both second and third trimesters of pregnancy. This classification relied on the measurement of the following: (1) Fetal abdominal circumference (AC), (2) Head circumference (HC), (3) BPD, and (4) Femur length.

Gestational age (GA) estimation

Classification and segmentation. Both classification and segmentation tasks were used to estimate GA in ($n = 3$, 2.80%) studies using 2D US taken in the second and third trimester of pregnancy. In (Maraci et al., 2020), the trans-cerebellar diameter (TCD) measurement was used to estimate GA in week. The TC plane frames are extracted from the US images using classification. Segmentation then localizes the TC structure and performs automated TCD estimation, from which the GA can thereafter be estimated via an equation. In (Chen et al., 2020a), 2D US was used to estimate GA based on region of fetal lung in the second and third trimester. In the first stage, the segmentation is to learn the recognition of fetal lung region in the ultrasound images. Classification is also used to accurately estimate the gestational age from the fetal lung region of ultrasound images. Several fetal structures were used to estimate GA in (Prieto et al., 2021), which focused on the second trimester. In this study, 2D US images were classified into four categories: head (BPD and HC), abdomen (AC), femur (FL), and fetus (crown-rump length: CRL). Then, the regions of interest (i.e., head, abdomen, and femur) were segmented and the results of biometry measurements were used to estimate GA.

Gender identification

Classification. Binary-classification was used in (Maysanjaya et al., 2014) to identify the gender of the fetus. Image preprocessing, image segmentation, and feature extraction (shape description) were used to obtain the value of the feature extraction and gender classification. This task is categorized as classification because a segmentation task is not clearly defined.

Fetus head

As shown in (Table 2), the primary purpose of ($n = 43$, 40.18%) studies is to identify and localize fetus skull (e.g., head circumference), brain standard planes (e.g., Lateral sulcus (LS), thalamus (T), choroid plexus (CP), cavum septi pellucidum (CSP)) or brain diseases (e.g., hydrocephalus (premature GA), ventriculomegaly, CNS malformations)). The following subsection discusses each category based on the implemented task. (Table 4) presents a comparison between the various techniques for each fetus head group (including objective, backbone methods, optimization, fetal age, best obtained result, and observations), studies for each group under the fetus head category, and provides an overview of methodology and observations.

Skull localization and measurement

Classification. Classification tasks for skull localization and HC are rarely used, as seen only in ($n = 3$, 2.80%) studies. Studies (Li et al., 2018; Nie et al., 2015a) used classification to localize the region of interest

Table 4. Articles published using AI to improve fetus head monitoring: objective, backbone methods, optimization, fetal age, and AI tasks

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|------------------------------------|---|---|---|----------------|----------------|
| Skull localization and measurement | | | | | |
| (Sobhaninia et al., 2020) | To localize the fetal head region in US imaging | Multi-scale mini-LinkNet network | N/A | 12 - 40 weeks | Segmentation |
| (Nie et al., 2015b) | To locate the fetal head from 3D ultrasound images using shape model | AdaBoost | Shape Model Marginal Space Haar-like features | 11 - 14 weeks | Classification |
| (Nie et al., 2015a) | To detect fetal head | Deep Belief Network (DBN) Restricted Boltzmann Machines | Hough transform Histogram Equalization | 11 - 14 weeks | Classification |
| (Aji et al., 2019) | To semantically segment fetal head from maternal and other fetal tissue | U-NET | Ellipse fitting | 12 - 20 weeks | Segmentation |
| (Droste et al., 2020) | To automatically discover and localize anatomical landmarks; measure the HC, TV, and the TC | CNN | Saliency maps | 13 - 26 weeks | Miscellaneous |
| (Desai et al., 2020) | To demonstrate the effectiveness of hybrid method to segment fetal head | DU-Net | Scattering Coefficients (SC) | 13 - 26 weeks | Segmentation |
| (Brahma et al., 2021) | To segment fetal head using Network Binarization | Depthwise Separable Convolutional Neural Networks DSCNNs. | Network Binarization | 12 - 40 weeks | Segmentation |
| (Qiao and Zulkernine, 2020) | To segment the fetal skull boundary and fetal skull for fetal HC measurement | U-NET | Squeeze and Excitation (SE) blocks | 12 - 40 weeks | Segmentation |
| (Sobhaninia et al., 2019) | To automatically segment and estimate HC ellipse. | Multi-Task network based on Link-Net architecture (MTLN) | Ellipse Tuner | 12 - 40 weeks | Segmentation |
| (Zhang et al., 2020b) | To capture more information with multiple-channel convolution from US images | Multiple-Channel and Atrous MA-Net | Encoder and Decoder Module | N/A | Segmentation |
| (Zeng et al., 2021) | To automatically segment fetal ultrasound image and HC biometry | Deeply Supervised Attention-Gated (DAG) V-Net | Attention-Gated Module | 12 - 40 weeks | Segmentation |
| (Perez-Gonzalez et al., 2020) | To compound a new US volume containing the whole brain anatomy | U-NET + Incidence Angle Maps (IAM) | CNN Normalized Mutual Information (NMI) | 13 to 26 weeks | Segmentation |

(Continued on next page)

Table 4. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|---|---|--|--|---------------|----------------|
| (Zhang et al., 2020a) | To directly measure the head circumference, without having to resort the handcrafted features or manually labeled segmented images. | CNN regressor (Reg-Resnet50) | N/A | 12 - 40 weeks | Segmentation |
| (Fiorentino et al., 2021) | To propose region-CNN for head localization and centering, and a regression CNN to accurately delineate the HC | CNN regressor (U-net) | Tiny-YOLOv2 | 12 - 40 weeks | Miscellaneous |
| (Li et al., 2020) | To present a novel end-to-end deep learning network to automatically measure the fetal HC, biparietal diameter (BPD), and occipitofrontal diameter (OFD) length from 2D US images | FCNN (SAPNet) | Regression network | 12 - 40 weeks | Miscellaneous |
| (Al-Bander et al., 2020) | To segment fetal head from US images | FCN | Faster R-CNN | 12 - 40 weeks | Miscellaneous |
| (Fathimuthu Joharah and Mohideen, 2020) | To deal with a completely computerized detection device of next fetal head composition | Multi-Task network based on Link-Net architecture (MTLN) | Hadamard Transform (HT) ANN FeedForward (NFFE) Classifier | N/A | Miscellaneous |
| (Li et al., 2018) | To measure HC automatically | Random Forest Classifier | Haar-like features ElliFit method | 18 - 33 weeks | Classification |
| (Sinclair et al., 2018) | To determine measurements of fetal HC and BPD | FNC | N/A | 18 - 22 weeks | Segmentation |
| (Yang et al., 2020b) | To segment the whole fetal head in US volumes | Hybrid attention scheme (HAS) | 3D U-NET + Encoder and Decoder architecture for dense labeling | 20 - 31 weeks | Segmentation |
| (Xu et al., 2021) | To segment fetal head using a flexibly plug-and-play module called vector self-attention layer (VSAL) | CNN | Vector Self-Attention Layer (VSAL) Context Aggregation Loss (CAL) | 12 - 40 weeks | Segmentation |
| (Cerrolaza et al., 2018) | To provide automatic framework for skull segmentation in fetal 3D US | Two-Stage Cascade CNN (2S-CNN) U-NET | Incidence Angle Map Shadow Casting Map | 20 - 36 weeks | Segmentation |

(Continued on next page)

Table 4. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|-----------------------------|--|---|--|--|-----------------------------|
| (Skeika et al., 2020) | To segment 2D ultrasound images of fetal skulls based on a V-Net architecture | Fully Convolutional Neural Network Combination (VNet-c) | N/A | 12 - 40 weeks | Segmentation |
| (Namburete and Noble, 2013) | To segment the cranial pixels in an ultrasound image using a random forest classifier | Random Forest Classifier | Simple Linear Iterative Clustering (SLIC) Haar Features | 25 - 34 weeks | Segmentation |
| (Budd et al., 2019) | To automatically estimate fetal HC | U-Net | Monte-Carlo Dropout | 18 - 22 weeks | Segmentation |
| Brain standard plane | | | | | |
| (Qu et al., 2020a) | To automatically recognize six standard planes of fetal brains. | CNN+ Transfer learning DCNN | N/A | 18 - 22 weeks 40 th week | Classification |
| (Cuingnet et al., 2013) | To help the clinician or sonographer obtain these planes of interest by finding the fetal head alignment in 3D US | Random forest classifier | Shape model and template deformation algorithm Hough transform | 19 - 24 weeks. | Classification segmentation |
| (Singh et al., 2021b) | To segment the fetal cerebellum from 2D US images | U-NET +ResNet (ResU-NET-C) | N/A | 18 - 20 weeks | Segmentation |
| (Yang et al., 2021b) | To detect multiple planes simultaneously in challenging 3D US datasets | Multi-Agent Reinforcement Learning (MARL) | RNN Neural Architecture Search (NAS) Gradient-based Differentiable Architecture Sampler (GDAS) | 19 - 31 weeks | Miscellaneous |
| (Lin et al., 2019b) | To detect standard plane and quality assessment | Multi-task learning Framework Faster Regional CNN (MF R-CNN) | N/A | 14 - 28 weeks | Miscellaneous |
| (Kim et al., 2019a) | To tackle the automated problem of fetal biometry measurement with a high degree of accuracy and reliability | U-Net, CNN | Bounding-box regression (object-detection) | N/A | Miscellaneous |
| (Lin et al., 2019a) | To determine the standard plane in US images | Faster R-CNN | Region Proposal Network (RPN) | 14 - 28 weeks | Miscellaneous |
| (Namburete et al., 2018) | To address the problem of 3D fetal brain localization, structural segmentation, and alignment to a referential coordinate system | Multi-Task FCN | Slice-Wise Classification | 18 - 34 weeks | Classification segmentation |

(Continued on next page)

Table 4. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|----------------------|---|--|--|----------------|-----------------------------|
| (Huang et al., 2018) | To simultaneously localize multiple brain structures in 3D fetal US | View-based Projection Networks (VP-Nets) | U-Net CNN | 20 - 29 weeks | Classification segmentation |
| (Qu et al., 2020b) | To automatically identify six fetal brain standard planes (FBSPs) from the non-standard planes. | Differential-CNN | Modified feature map | 16 - 34 weeks | Classification |
| (Wang et al., 2019) | To obtain the desired position of the gate and Middle Cerebral Artery (MCA) | MCANet | Dilated Residual Network (DRN) Dense Upsampling Convolution (DUC) block | 28 - 40 weeks | Segmentation |
| (Yaqub et al., 2013) | To segment four important fetal brain structures in 3D US | Random Decision Forests (RDF) | Generalized Haar-features | 18 - 26 weeks | Segmentation |
| (Yang et al., 2021a) | To automatically localize fetal brain standard planes in 3D US | Dueling Deep Q Networks (DDQN) | RNN-based Active Termination (AT) (LSTM) | 19 - 31 weeks | Miscellaneous |
| (Liu et al., 2020) | To evaluate the feasibility of CNN-based DL algorithms predicting the fetal lateral ventricular width from prenatal US images. | ResNet50 | Faster R-CNN Class Activation Mapping (CAM) | 22 - 26 weeks. | Miscellaneous |
| (Sahli et al., 2020) | To recognize and separate the studied US data into two categories: healthy (HL) and hydrocephalus (HD) subjects | CNN | N/A | 20 - 22 weeks. | Classification |
| (Chen et al., 2020b) | To automatically measure fetal lateral ventricles (LVs) in 2D US images | Mask R-CNN | Feature Pyramid Networks (FPN) Region Proposal Network (RPN) | N/A | Miscellaneous |
| (Xie et al., 2020b) | To apply binary classification for central nervous system (CNS) malformations in standard fetal US brain images in axial planes | CNN | Split-view Segmentation | 18 - 32 weeks | Classification segmentation |
| (Xie et al., 2020a) | To develop computer-aided diagnosis algorithms for five common fetal brain abnormalities. | Deep convolutional neural networks (DCNNs) VGG-net | U-net Gradient-Weighted Class Activation Mapping (Grad-CAM) | 18 - 32 weeks | Classification segmentation |

(ROI) and identify the fetus head based on 2D US taken in the first and second trimesters respectively. 3D US taken in the first trimester were used to detect fetal head in one study (Nie et al., 2015b).

Segmentation. Skull localization and HC in most of the studies ($n = 16$, 14.81%) used segmentation task. 2D US was used in 16 studies and one study used 3D US. Various network architecture were used to segment and locate the skull and perform HC as seen in (Aji et al., 2019; Brahma et al., 2021; Budd et al., 2019; Desai et al., 2020; Namburete and Noble, 2013; Perez-Gonzalez et al., 2020; Qiao and Zulker-nine, 2020; Skeika et al., 2020; Sobhaninia et al., 2020,2019; Xu et al., 2021; Zeng et al., 2021; Zhang et al., 2020b). Besides identifying HC, in (Sinclair et al., 2018) segmentation was also used to find fetal BPD. Further investigation shows that the work completed in (Cerrolaza et al., 2018) was the first investigation about whole fetal head segmentation in 3D US. The segmentation network for skull localization in (Xu et al., 2021) was tested to identify a view of the four heart chambers on different datasets.

Miscellaneous. In addition to segmentation, another method was used to locate and identify the fetus skull in ($n = 6$, 5.60%) studies. As seen in (Li et al., 2020), regression was used with segmentation to identify HC, BPD, and occipitofrontal diameter (OFD). Besides this, in (Droste et al., 2020), neural network was used to train the saliency predictor and predict saliency maps to measure HC and identify the alignment of both transventricular (TV) as well as transcerebellar (TC). In addition, an object detection task was used with segmentation in (Al-Bander et al., 2020) to locate the fetal head at the end of the first trimester. In (Fiorentino et al., 2021), object detection was used for head localization and centering and regression to delineate the HC accurately. In (Fathimuthu Joharah and Mohideen, 2020), classification and segmentation tasks were used by utilizing a multi-task network to identify head composition.

Brain standard plane

Classification. The classification was used in ($n = 2$, 1.86%) studies to identify the brain standard plane, and both studies used 2D US images taken in the second and third trimesters. In (Qu et al., 2020a; 2020b), two different classification network architectures were used to identify six fetal brain standard planes. These architectures include horizontal transverse section of the thalamus, horizontal transverse section of the lateral ventricle, transverse section of the cerebellum, mid-sagittal plane, paracentral sagittal section, and coronal section of the anterior horn of the lateral ventricles.

Segmentation. The segmentation task was used in ($n = 3$, 2.80%) studies to identify the brain standard plane. Each of these studies used 2D US images. In (Singh et al., 2021b), 2D US images taken in the second trimester were used to segment fetal cerebellum structures. In addition, in (Wang et al., 2019), 2D US images taken in the third trimester were used to segment fetal middle cerebral artery (MCA). Authors in (Yaqub et al., 2013) utilized 3D US images in the second trimester to formulate the segmentation as a classification problem to identify the following brain planes; background, choroid plexus (CP), lateral posterior ventricle cavity (PVC), the cavum septum pellucidi (CSP), and cerebellum (CER).

Classification and segmentation. Classification with segmentation is employed in ($n = 3$, 2.80%) studies to identify the brain standard plane. These studies all used 3D US images. In (Namburete et al., 2018), both tasks were used to identify brain alignment based on skull boundaries and then head segmentation, eye localization, and prediction of brain orientation in the second and third trimesters. Moreover, in (Huang et al., 2018), segmentation followed by a classification task was used to detect CSP, Tha, lateral ventricles (LV), cerebellum (CE), and cisterna magna (CM) in the second trimester. Segmentation and classification are employed by 3D US in (Cuingnet et al., 2013) to identify the skull, mid-sagittal plane, and orbits of the eyes in the second trimester of pregnancy.

Miscellaneous. For brain standard planes identification in ($n = 5$, 4.62%), methods different than previously mentioned were used, including; segmentation with object detection, object detection, classification with object detection, and RL. As seen in (Yang et al., 2021a), the RL-based technique was employed for the first time on 3D US images to localize standard brain planes, including *trans*-thalamic (TT) and *trans*-cerebellar (TC) in the second and third trimesters. RL is used in (Yang et al., 2021b) to localize the following: mid-sagittal (S), transverse (T), and coronal (C) planes in volumes and *trans*-thalamic (TT), *trans*-ventricular (TV), and *trans*-cerebellar (TC)). Object detection architecture is utilized on 2D US images in (Lin et al., 2019a) to localize the *trans*-thalamic plane in the second trimester of pregnancy, including LS, T, CP, CSP, and third ventricle (TV). In (Lin et al., 2019b), classification with object detection was used to detect

LS, T, CP, CSP, TV, brain midline (BM). In contrast (Kim et al., 2019a), used classification with object detection to localize cavum septum pellucidum (CSP) and ambient cistern (AC) and cerebellum, as well as to measure HC and BPD.

Brain disease

Classification. The binary classification technique utilized 2D US images taken in the second trimester as seen in (Sahli et al., 2020) to detect hydrocephalus disease using premature GA.

Classification and segmentation. Studies (Xie et al., 2020a, 2020b) both used 2D US images taken in the second and third trimesters to segment craniocerebral and identify abnormalities or specific diseases. As seen in (Xie et al., 2020b), CNS malformations were detected using binary classification. Moreover, in (Xie et al., 2020a), multi-classification was used to identify the following problems: TV planes contained occurrences of ventriculomegaly and hydrocephalus, and TC planes contained occurrences of Blake pouch cyst (BPC).

Miscellaneous. In (Chen et al., 2020b; Liu et al., 2020) different methods were used to detect ventriculomegaly disease based on 2D US images taken in the second trimester. In study (R. Liu et al., 2020), object detection and regression first identify fetal brain ultrasound images from standard axial planes as normal or abnormal. Second regression was then used to find the lateral ventricular regions of images with big lateral ventricular width. Then, the width was anticipated with a modest error based on these regions. Furthermore (Chen et al., 2020b), was the first study to propose an object-detection-based automatic measurement approach for fetal lateral ventricles (LVs) based on 2D US images. The approach can both distinguish and locate the fetal LV automatically as well as measure the LV's scope quickly and accurately.

Fetus face

As shown in (Table 2), the primary purpose of (n= 10, 9.34%) studies was to identify and localize fetus face features such as the fetal facial standard plane (FFSP) (i.e., axial, coronal, and sagittal plane), face anatomical landmarks (e.g., nasal bone), and facial expressions (i.e., sad, happy). The following subsection discusses each category based on the implemented task. (Table 5). presents comparisons between the various techniques for each fetus face group including objective, backbone methods, optimization, fetal age, best obtained result, and observations.

Fetal facial standard plane (FFSP)

Classification. Classification was used to classify the FFSP using 2D US Images as seen in (n = 5, 4.67%). Using 2D US images that were taken in the second and third trimesters, four studies (Lei et al., 2014, 2015; Yu et al., 2016, 2018) identify ocular axial planes (OAP), the median sagittal planes (MSP), and the nasolabial coronal planes (NCP). In addition, authors in (Wang et al., 2021) were able to identify the FFSP using 2D US images taken in the second trimester.

Face anatomical landmarks

Miscellaneous. Different methods were used to identify face anatomical landmarks in (n = 3, 2.80) studies. In (Singh et al., 2021a), 3D US images taken in the second and third trimesters were segmented to identify background, face mask (excluding facial structures), eyes, nose, and lips. In another study the object detection method was used on 3D US images in (Chen et al., 2020c) to detect the left fetal eye, middle eyebrow, right eye, nose, and chin. Furthermore, classification used 2D US images taken in the first and second trimesters to detect the nasal bone. This was done to enhance the detection rate of Down syndrome, as seen in (Anjit and Rishidas, 2011).

Facial expressions

Classification. For the first time, 4D US images were utilized using the multi-classification method to identify fetus facial expression into Sad, Normal, and Happy as seen in (Dave and Nadiad, 2015). Study (Miyagi et al., 2021) used 2D US images in the second and third trimesters to classify fetal facial expression into eye blinking, mouthing without any expression, scowling, and yawning.

Fetus heart

As shown in (Table 2), the primary purpose of (n= 13, 14.04%) studies is to identify and localize fetus heart diseases and the fetus heart chambers view. The following subsection discusses each category based on the implemented task. (Table 6). presents comparisons between the various techniques for each fetus heart

Table 5. Articles published using AI to improve fetus face monitoring: objective, backbone methods, optimization, fetal age, and AI tasks

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|---|---|---|---|---------------|----------------|
| Fetal facial standard plane (FFSP) | | | | | |
| (Lei et al., 2014) | To address the issue of recognition of standard planes (i.e., axial, coronal and sagittal planes) in the fetal US image | SVM classifier | AdaBoost for detect region of interest, ROI Dense Scale Invariant Feature Transform (DSIFT) Aggregating vectors for feature extraction fish vector (FV) Gaussian Mixture Model (GMM) | 20 - 36 weeks | Classification |
| (Yu et al., 2016) | To automatically recognize the FFSP from US images | Deep convolutional networks (DCNN) | N/A | 20 - 36 weeks | Classification |
| (Yu et al., 2018) | To automatically recognize FFSP via a deep convolutional neural network (DCNN) architecture | DCNN | t-Distributed Stochastic Neighbor Embedding (t-SNE) | 20 - 36 weeks | Classification |
| (Lei et al., 2015) | To automatically recognize the fetal facial standard planes (FFSPs) | SVM classifier | Root scale invariant feature transform (RootSIFT) Gaussian mixture model (GMM) Fisher Vector (FV) Principal Component Analysis (PCA) | 20 - 36 weeks | Classification |
| (Wang et al., 2021) | To automatically recognize and classify FFSPs | SVM classifier | Local Binary Pattern (LBP) Histogram of Oriented Gradient (HOG) | 20 - 24 weeks | Classification |
| Face anatomical landmarks | | | | | |
| (Singh et al., 2021a) | To detect position and orientation of facial region and landmarks | SFFD-Net (Samsung Fetal Face Detection Network) multi-class segmented | N/A | 14 - 30 weeks | Miscellaneous |
| (Chen et al., 2020c) | To detect landmarks in 3D fetal facial US volumes | CNN Backbone Network | Region Proposal Network (RPN) Bounding-box regression | N/A | Miscellaneous |
| (Anjit and Rishidas, 2011) | To detect nasal bone for US of fetus | Back Propagation Neural Network (BPNN) | Discrete Cosine Transform (DCT) Daubechies D4 Wavelet transform | 11 - 13 weeks | Miscellaneous |

(Continued on next page)

Table 5. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|-------------------------|--|--------------------------------|--|---------------|----------------|
| Facial expressions | | | | | |
| (Dave and Nadiad, 2015) | To recognize facial expressions from 3D US | ANN | Histogram equalization Thresholding Morphing Sampling Clustering Local Binary Pattern (LBP) Minimum Redundancy and Maximum Relevance (MRMR) | N/A | Classification |
| (Miyagi et al., 2021) | To recognize fetal facial expressions that are considered as being related to the brain development of fetuses | CNN | N/A | 19 - 38 weeks | Classification |

Table 6. Articles published using AI to improve fetus heart monitoring: Objective, backbone methods, optimization, fetal age, and AI tasks

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|----------------------------------|--|---|---|-------------|--------------------------------|
| Heart disease | | | | | |
| (Yang et al., 2020a) | To perform multi-disease segmentation and multi-class semantic segmentation of the five key components | U-NET + DeepLabV3+ | N/A | N/A | Segmentation |
| (Gong et al., 2020) | To recognize and judge fetal congenital heart disease (FHD) development | DGACNN Framework | CNN Wasserstein GAN + Gradient Penalty (WGAN-GP) DANomaly Faster-RCNN | 18–39 weeks | Miscellaneous |
| (Dozen et al., 2020) | To segment the ventricular septum in US | Cropping-Segmentation-Calibration (CSC) | YOLOv2 cropping module U-NET segmentation Module VGG-backbone Calibration Module | 18-28 weeks | Miscellaneous |
| (Komatsu et al., 2021b) | To detect cardiac substructures and structural abnormalities in fetal US videos | Supervised Object detection with Normaldata Only (SONO) | CNN YOLOv2 | 18-34 weeks | Miscellaneous |
| (Arnaout et al., 2021) | To identify recommended cardiac views and distinguish between normal hearts and complex CHD and to calculate standard fetal cardiothoracic measurements | Ensemble of Neural Networks | ResNet and U-Net Grad-CAM | 18-24 weeks | Classification segmentation |
| (Dinesh Simon and Kavitha, 2021) | To learn the features of Echogenic Intracardiac Focus (EIF) that can cause Down Syndrome (DS) whereas testing phase classifies the EIF into DS positive or DS negative based | Multi-scale Quantized Convolution Neural Network (MSQCNN) | Cross-Correlation Technique (CCT) Enhanced Learning Vector Quantiser (ELVQ) | 24–26 weeks | Classification |

(Continued on next page)

Table 6. Continued

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|---------------------------|--|--------------------------------|--|---------------|----------------|
| (Tan et al., 2020) | To perform automated diagnosis of hypoplastic left heart syndrome (HLHS) | SonoNet (VGG16) | N/A | 18–22 weeks | Classification |
| Heart chamber's view | | | | | |
| (Xu et al., 2020b) | To perform automated segmentation of cardiac structures | CU-NET | Structural Similarity Index Measure (SSIM) | N/A | Segmentation |
| (Xu et al., 2020a) | To accurately segment seven important anatomical structures in the A4C view | DW-Net | Dilated Convolutional Chain (DCC) module W-Net module based on the concept of stacked U-Net | N/A | Segmentation |
| (Dong et al., 2020) | To automatically quality control the fetal US cardiac four-chamber plane | Three CNN-based Framework | Basic-CNN, a variant of SqueezeNet Deep-CNN with DenseNet-161 as basic architecture The ARVNet for real-time object detection. | 14 - 28 weeks | Miscellaneous |
| (Pu et al., 2021) | To localize the end-systolic (ES) and end-diastolic (ED) from ultrasound | Hybrid CNN based framework | YOLOv3 Maximum Difference Fusion (MDF) Transferred CNN | 18 - 36 weeks | Miscellaneous |
| (Sundaresan et al., 2017) | To detect the fetal heart and classifying each individual frame as belonging to one of the standard viewing planes | FCN | N/A | 20 - 35 weeks | Segmentation |
| (Patra et al., 2017) | To jointly predict the visibility, view plane, location of the fetal heart in US videos. | Multi-Task CNN | Hierarchical Temporal Encoding (HTE) | 20 - 35 weeks | Classification |

group including objective, backbone methods, optimization, fetal age, best obtained result, and observations.

Heart disease

Classification. Two studies used classification methods to identify heart diseases based on 2D US images taken in the second trimester. In (Dinesh Simon and Kavitha, 2021), binary-classification task was utilized to identify Down syndrome versus normal fetuses based on identifying the echogenic intracardiac foci (EIF). Furthermore, in (Tan et al., 2020), binary-classification task was utilized to detect hypoplastic left heart syndrome (HLHS) versus healthy cases based on the four-chamber heart (4CH), left ventricular outflow tract (LVOT), and right ventricular outflow tract (RVOT).

Segmentation. 2D US images were utilized by multi-class segmentation to identify heart disease, including: left heart syndrome (HLHS), total anomalous pulmonary venous connection (TAPVC), pulmonary atresia with intact ventricular septum (PA/IVS), endocardial cushion defect (ECD), fetal cardiac rhabdomyoma (FCR), and Ebstein's anomaly (EA) (Yang et al., 2020a).

Classification and segmentation. Classification and segmentation are used to detect congenital heart disease (CHD) based on 2D US images taken in the second trimester (Arnaout et al., 2021). Cases were classified into normal hearts vs CHD. This classification was orchestrated by identifying five views of the heart in fetal CHD screening, including three-vessel trachea (3VT), three-vessel view (3VV), left-ventricular outflow tract (LVOT), axial four chambers (A4C), and abdomen (ABDO).

Miscellaneous. To identify heart disease, various methods were proposed based on the capabilities of 2D US images. Classification with object detection was utilized in (Gong et al., 2020; Komatsu et al., 2021b). Fetal congenital heart disease (FHD) can be detected in the second and third trimesters based on how quickly the fetus grows between gestational weeks, and how the shape of the four chambers in the heart changes over time (Gong et al., 2020). Furthermore, classification with object detection was used to detect cardiac substructures and structural abnormalities in the second and third trimesters (Komatsu et al., 2021b). Lastly, segmentation with object detection was used to locate the ventricular septum in 2D US images in the second trimester, as seen in (Dozen et al., 2020).

Heart chambers view

Classification. 2D US images taken in the second and third trimesters were utilized in (Patra et al., 2017) to propose a classification task utilized to localize the four chambers (4C), the left ventricular outflow tract (LVOT), the three vessels (3V), and the background (BG).

Segmentation. Segmentation was used in ($n = 3$, 2.80%) studies based on 2D US images. In two studies (Xu et al., 2020a, 2020b), seven critical anatomical structures in the apical four-chamber (A4C) view were segmented, including: left atrium (LA), right atrium (RA), left ventricle (LV), right ventricle (RV), descending aorta (DAO), epicardium (EP) and thorax. In another study (Sundaresan et al., 2017), segmentation was used to locate four-chamber views (4C), left ventricular outflow tract view (LVOT), and three-vessel view (3V) in the second and third trimesters.

Miscellaneous. 2D US images were used in ($n = 2$, 1.86) studies, and both studies utilized classification with object detection. In the second trimester (Dong et al., 2020), the first classification conducted was the cardiac four-chamber plane (CFP) into non-CFPs and CFPs (i.e., apical, bottom, and parasternal CFPs). The second task was to then classify the CFPs in terms of the zoom and gain of 2D US images. In addition, object detection was utilized to detect anatomical structures in the CFPs, including left atrial pulmonary vein angle (PVA), apex cordis and moderator band (ACMB), and multiple ribs (MRs). In (Pu et al., 2021), object detection was used in the second and third trimesters to extract attention regions for improving classification performance and determining the four-chamber view, including detection of end-systolic (ES) and end-diastolic (ED).

Fetus abdomen

As shown in (Table 2), the primary purpose of ($n = 10$, 10.8%) studies was to identify and localize the fetus's abdomen, including abdominal anatomical landmarks (i.e., stomach bubble (SB), umbilical vein (UV), and spine (SP)). The following subsection discusses this category based on the implemented task. (Table 7).

Table 7. Articles published using AI to improve fetus abdomen monitoring: objective, backbone methods, optimization, fetal age, and AI tasks

| Study | Objective | Backbone Methods/ Framework | Optimization/Extractor methods | Fetal age | AI tasks |
|---------------------------------------|---|---|--|----------------|-----------------------------|
| Abdominal anatomical landmarks | | | | | |
| (Rahmatullah et al., 2011b) | To automatically detect two anatomical landmarks in an abdominal image plane stomach bubble (SB) and the umbilical vein (UV). | AdaBoost | Haar-like feature | 14 - 19 weeks | Classification |
| (Yang et al., 2014) | To localize fetal abdominal standard plane (FASP) from US including SB, UV, and spine (SP) | Random Forests Classifier+ SVM | Haar-like feature Radial Component-Based Model (RCM) | 18 - 40 weeks | Classification |
| (Kim et al., 2018) | To classify ultrasound images (SB, amniotic fluid (AF), and UV) and to obtain an initial estimate of the AC." | Initial Estimation CNN + U-Net | Hough transform | N/A | Classification segmentation |
| (Jang et al., 2017) | To classify ultrasound images (SB, AF, and UV) and measure AC | CNN | Hough transform | 20 - 34 weeks | Classification segmentation |
| (Wu et al., 2017) | To find the region of interest (ROI) of the fetal abdominal region in the US image. | Fetal US Image Quality Assessment (FUIQA) | L-CNN is able to localize the fetal abdominal ROI AlexNet C-CNN then further analyzes the identified ROI DCNN to duplicate the US images for the RGB channels rotating" | 16 - 40 weeks | Classification |
| (Ni et al., 2014) | To localize the fetal abdominal standard plane from ultrasound | Random forest classifier+ SVM classifier | Radial Component-based Model (RCM) Vessel Probability Map (VPM) Haar-like features | 18 - 40 weeks | Classification |
| (Deepika et al., 2021) | To diagnose the (prenatal) US images by design and implement a novel framework | Defending Against Child Death (DACD) | CNN U-Net Hough-man transformation | N/A | Classification segmentation |
| (Rahmatnllah et al., 2012) | To detect important landmarks employed in manual scoring of ultrasoundimages. | AdaBoost | Haar-like feature | 18 - 37 weeks | Classification |
| (Rahmatullah et al., 2011a) | To automatically select the standard plane from the fetal US volume for the application of fetal biometry measurement. | AdaBoost | One Combined Trained Classifier (1CTC) Two Separately Trained Classifiers (2STC) Haar-like feature | 20 - 28 weeks | Classification |
| (Chen et al., 2014) | To localize the FASP from US images. | DCNN | Fine-Tuning with Knowledge Transfer Barnes-Hut Stochastic Neighbor Embedding (BH-SNE) | 18 - 40 weeks) | Classification |

presents comparisons between the various techniques for each fetus abdomen including objective, back-bone methods, optimization, fetal age, best obtained result, and observations.

Abdominal anatomical landmarks

Classification. Classification task was used in ($n = 7$, 6.54%) studies to localize abdominal anatomical landmarks. 2D US images were used in seven studies (Chen et al., 2014; Ni et al., 2014; Rahmatnillah et al., 2012; Rahmatullah et al., 2011b; Wu et al., 2017; Yang et al., 2014) and 3D US images in only one study (Rahmatullah et al., 2011a). In (Rahmatnillah et al., 2012; Rahmatullah et al., 2011a, 2011b; Wu et al., 2017), various classifiers were utilized to localize stomach bubble (SB) and umbilical vein (UV), in the first and second trimester as seen in (Rahmatullah et al., 2011a, 2011b), and in the second and third trimesters as seen in (Rahmatnillah et al., 2012; Wu et al., 2017). Furthermore, works in (Chen et al., 2014; Ni et al., 2014; Yang et al., 2014) located the spine (SP) besides SB and UV on the same trimesters.

Classification and segmentation. 2D US images taken within random trimesters were used in (Deepika et al., 2021; Kim et al., 2018). In (Kim et al., 2018), SB, UV, and AF were localized. Further, abdominal circumference (AC), spine position, and bone regions were estimated. In (Deepika et al., 2021), 2D US images were classified into normal versus abnormal fetuses based on the fetus images such as AF, SB, UV, and SA. Study (Jang et al., 2017) located SB and the portal section from the UV, and observed amniotic fluid (AF) in the second and third trimesters.

Dataset analysis

Fetus body

In this area, ethical and legal concerns presented the most barriers to comprehensive research. Therefore, datasets were not available to the research community. For example, in this research, studies discussed how the fetal body contains five subsections (fetal part structure, anatomical structure, growth disease, gestational age, and gender identification). Of the available studies ($n = 31$, 28.9%), one dataset was intended to be available online in (Yaqub et al., 2015) related to identifying fetal part structure. Unfortunately, this dataset has not yet been released by the author. The only public dataset that was available online for fetal structure classification was released by Burgos-Artizzu et al. (2020). The dataset is a total of 12,499 2D fetal ultrasound images, including brain 143, Trans-cerebellum 714, Trans-thalamic 1638, Trans-ventricular 597, abdomen 711, cervix 1626, femur 1040, thorax 1718, and 4213 are unclassified fetal images. Acquiring a high volume of the dataset was also challenging in most of the studies; therefore, multi-data augmentation in ($n = 11$, 10.2%) studies (Andriani and Mardhiyah, 2019; Burgos-Artizzu et al., 2020; Cai et al., 2018; Gao and Noble, 2019; Li et al., 2017; Maraci et al., 2020; Ryou et al., 2019; Weerasinghe et al., 2021; Yang et al., 2019) has been observed. These augmented images were employed to boost classification and segmentation performance. SimpleITK library (Lowekamp et al., 2013) was used for augmentation in (Weerasinghe et al., 2021). Because of limited data samples in addition to augmentation, k-fold cross-validation method was utilized in ($n = 8$, 7.4%) studies (Bagi and Shreedhara, 2014; Cai et al., 2020; Gao and Noble, 2019; Maysanjaya et al., 2014; Rahmatullah et al., 2014; Ryou et al., 2016; Wee et al., 2010; Yaqub et al., 2015). These cross-validation methods are also used to resolve issues such as overfitting and bias in dataset selection. Only one software named ITK-SNAP (Yushkevich et al., 2006) was reported in (Weerasinghe et al., 2021) to segment structures in 3D images.

Fetus head

This section contains three subsections (skull localization, brain standard planes, and brain disease). There were a total of ($n = 43$, 40.18%) studies, but we only found one online public dataset called HC18 grand challenge (van den Heuvel et al., 2018). This dataset was used in ($n = 13$, 12.14%) studies (Aji et al., 2019; Al-Bander et al., 2020; Brahma et al., 2021; Desai et al., 2020; Fiorentino et al., 2021; Li et al., 2020; Qiao and Zulkernine, 2020; Skeika et al., 2020; Sobhaninia et al., 2020, 2019; Xu et al., 2021; Zeng et al., 2021; Zhang et al., 2020a). This dataset contains 2D US images collected from 551 pregnant women at different pregnancy trimesters. There were 999 images in the training set and 335 for testing; the sonographer manually annotated the HC. Image augmentation was applied in ($n = 9$, 8.41%) studies (Al-Bander et al., 2020; Brahma et al., 2021; Fiorentino et al., 2021; Li et al., 2020; Skeika et al., 2020; Sobhaninia et al., 2019, 2020; Zeng et al., 2021; Zhang et al., 2020a), and cross-validation was employed in (Zhang et al., 2020a).

(Table 8) highlights the best result achieved despite the challenges of recording the outcomes of the selected studies.

Table 8. Compared between studies that utilized the HC18 dataset

| Study | DSC | HD | DF | ADF |
|----------------------------|--------------|------|-------------|-------------|
| (Sobhaninia et al., 2020) | 0.926 | 3.53 | 0.94 | 2.39 |
| (Aji et al., 2019) | N/A | N/A | 14.9% | N/A |
| (Desai et al., 2020) | 0.973 | 1.58 | N/A | N/A |
| (Brahma et al., 2021) | 0.968 | N/A | N/A | N/A |
| (Qiao and Zulkemine, 2020) | 0.973 | N/A | N/A | 2.69 |
| (Sobhaninia et al., 2019) | 0.968 | 1.72 | 1.13 | 2.12 |
| (Zeng et al., 2021) | 0.979 | 1.27 | 0.09 | 1.77 |
| (Fiorentino et al., 2021) | 0.977 | 1.32 | 0.21 | 1.90 |
| (Li et al., 2020) | 0.977 | 0.47 | N/A | 2.03 |
| (Al-Bander et al., 2020) | 0.977 | 1.39 | 1.49 | 2.33 |
| (Xu et al., 2021) | 0.971 | 3.23 | N/A | N/A |
| (Skeika et al., 2020) | 0.979 | N/A | N/A | N/A |

DSC, Dice similarity coefficient; ACC, Accuracy; Pre, Precision; HD, Hausdorff distance; DF, Difference; ADF, Absolute Difference; IoU, Intersection overUnion; mPA, mean Pixel Accuracy.

Fetus face

This section contains three subsections (fetal facial standard planes, face anatomical landmarks, and facial expression). Of these studies (n = 10, 9.34%), we did not find any public dataset available online. However, within the private dataset, image augmentation was applied in (n = 3, 2.8%) studies (Chen et al., 2020c; Miyagi et al., 2021; Singh et al., 2021a). The k-fold cross-validation was employed in (Lei et al., 2014; Singh et al., 2021a; Wang et al., 2021). Lastly, 3D Slicer software (Pieper et al., 2004) was used for image annotation as reported in (Singh et al., 2021a).

Fetus heart

This section contains two subsections (heart diseases and heart chamber view). Of these studies (n = 13, 12.14%), we did not find any public dataset available online. However, within the private dataset, the image augmentation was applied in (n = 4, 3.7%) studies (Arnaout et al., 2021; Pu et al., 2021; Sundaresan et al., 2017; Yang et al., 2020a). The k-fold cross-validation was employed in (Arnaout et al., 2021; Dong et al., 2020; Dozen et al., 2020; Gong et al., 2020; Xu et al., 2020a).

Fetus abdomen

This section contains one subsection (abdominal anatomical landmarks). Of these studies (n = 10, 9.34%), we did not find any public dataset available online. However, within the private dataset, image augmentation was applied in (n = 4, 3.7%) studies (Chen et al., 2014; Jang et al., 2017; Kim et al., 2018; Wu et al., 2017). The cross-validation was not reported in any of the studies.

DISCUSSION

Principal findings

AI techniques applied are shown in Figure 4. Deep learning (DL) was the most utilized techniques, seen in (n = 81, 75.70%) studies. Traditional Machine learning (ML) models are used in (n = 16, 14.95%) studies. Artificial neural network (ANN) models are only used in (n = 7, 6.5%) studies. We also found the use of RL in (n = 2, 1.86%) studies. Both deep learning and machine learning models were used in one study.

Deep learning for fetus health

As seen in Figure 4. DL was used heavily in all fetal organ experiments including general fetus issues, fetus head, fetus face, fetus heart, and fetus abdomen. Furthermore, as seen in most of the models (n = 74, 69.1%), Convolutional Neural Network (CNN) used as a backbone of deep learning and is used in all tasks including classification, segmentation, and object detection (Aji et al., 2019; Al-Bander et al., 2020; Andriani and Mardhiyah, 2019; Arnaout et al., 2021; Brahma et al., 2021; Budd et al., 2019; Cai et al., 2020, 2018; Chen et al., 2014, 2017, 2020a, 2020b, 2020c; Deepika et al., 2021; Desai et al., 2020; Dinesh Simon and Kavitha, 2021; Dong et al., 2020; Dozen et al., 2020; Droste et al., 2020; Fathimathu Joharah and Mohideen, 2020;

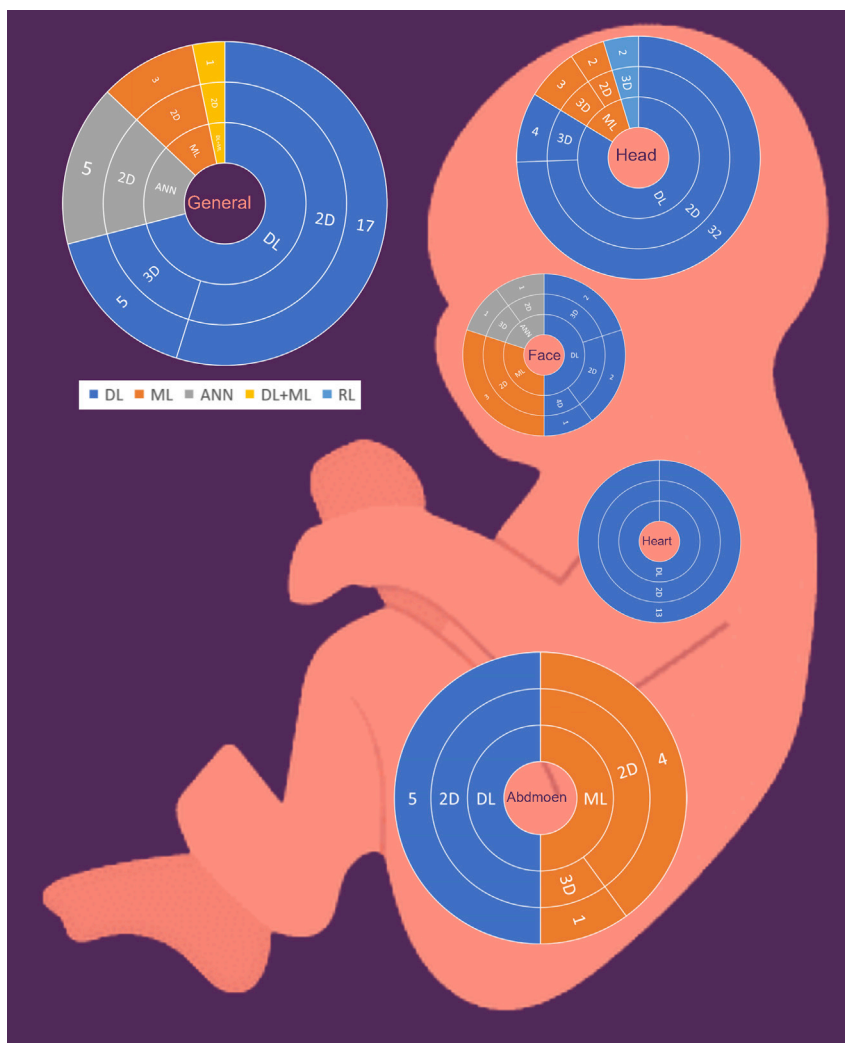


Figure 4. AI techniques used within all studies in this survey, AI techniques are ordered by US images type and further categorized by fetal organ. Totals here equal the number of included papers

Fiorentino et al., 2021; Gao and Noble, 2019; Gong et al., 2020; Huang et al., 2018; Jang et al., 2017; Kim et al., 2018, 2019a; Komatsu et al., 2021b; Li et al., 2020; Li et al., 2017; Lin et al., 2019a, 2019; Liu et al., 2020; Liu et al., 2019; Miyagi et al., 2021; Namburete and Noble, 2013; Nie et al., 2015a; Patra et al., 2017; Perez-Gonzalez et al., 2020; Prieto et al., 2021; Pu et al., 2021; Qiao and Zulkernine, 2020; Qu et al., 2020a; Ravishankar et al., 2016; Ryou et al., 2016, 2019; Sahli et al., 2020; Singh et al., 2021a, 2021b; Sobhaninia et al., 2019, 2020; Tan et al., 2020; Toussaint et al., 2018; Wang et al., 2019; Weerasinghe et al., 2021; Wu et al., 2017; Xie et al., 2020a, 2020b; Xu et al., 2021; Xu et al., 2020a, 2020b; Yang et al., 2020a, 2020b; Yekdast, 2019; Yu et al., 2016, 2018; Zeng et al., 2021; Zhang et al., 2020a, 2020b). In addition, Fully Connected Neural Networks (FCNNs) are similarly largely utilized as a backbone or part of the framework ($n = 11$, 10.28%) of DL, as seen in (Al-Bander et al., 2020; Dong et al., 2020; Looney et al., 2021; Maraci et al., 2020; Namburete et al., 2018; Ryou et al., 2019; Sinclair et al., 2018; Skeika et al., 2020; Sundaresan et al., 2017; Weerasinghe et al., 2021; Yang et al., 2019). In addition to CNN and FCN, a Recurrent Neural Network (RNN) used to perform classification tasks in hybrid frameworks was seen in ($n = 5$, 4.67%) studies (Cai et al., 2020; Chen et al., 2015, 2017; Gao and Noble, 2019; Yang et al., 2019). U-Net is a convolutional neural network that was developed for biomedical image segmentation. Therefore, we found U-NET to be the most frequently utilized segmentation model as a baseline of the framework or as optimization. This was seen in ($n = 23$, 21.50%) (Aji et al., 2019; Arnaout et al., 2021; Budd et al., 2019; Cerrolaza et al., 2018; Chen et al., 2020a; Deepika et al., 2021; Desai et al., 2020; Dozen et al., 2020; Huang et al., 2018; Kim et al.,

Table 9. Best DL study in each category based on the achieved result

| Mian Organ | Subsection | Best study |
|------------|-----------------------|---|
| Fetal Body | Fetal part structures | (Liu et al., 2021, 2019) |
| | Anatomical structures | (Liu et al., 2019; Toussaint et al., 2018) |
| | Growth disease | (Yekdast, 2019) |
| | Gestational age | (Chen et al., 2020a) |
| Head | Skull localization | (Fiorentino et al., 2021; Zhang et al., 2020a) |
| | Brain Standard plan | (Qu et al., 2020a,2020b; Wang et al., 2019) |
| | Brain disease | (Chen et al., 2020b; Sahli et al., 2020) |
| Face | Fetal facial standard | (Yu et al., 2018) |
| | Face anatomical | (Singh et al., 2021a) |
| | Facial expression | (Miyagi et al., 2021) |
| Heart | Heart disease | (Arnaout et al., 2021; Dinesh Simon and Kavitha, 2021; Komatsu et al., 2021b; Yang et al., 2020a) |
| | Hear chamber view | (Pu et al., 2021; Xu et al., 2020a) |
| Abdomen | Abdominal anatomical | (Deepika et al., 2021; Kim et al., 2018) |

2018, 2019a; Lin et al., 2019a; Liu et al., 2019; Perez-Gonzalez et al., 2020; Prieto et al., 2021; Qiao and Zulkernine, 2020; Singh et al., 2021b; Weerasinghe et al., 2021; Xie et al., 2020a; Xu et al., 2020a, 2020b; Yang et al., 2020a; Yang et al., 2020b). Other segmentation models were also used, such as DeepLabV3 in (Yang et al., 2020a), LinkNet in (Fathimithu Joharah and Mohideen, 2020), and the Encoder-Decoder network based on VGG16 in (Li et al., 2017). A Residual Neural Network (ResNet) was used to perform some tasks and enhance the framework efficiency as seen in (n = 6, 5.60%) studies (Arnaout et al., 2021; Liu et al., 2020, 2021; Prieto et al., 2021; Singh et al., 2021b; Toussaint et al., 2018; Wang et al., 2019). Some popular object detection models (n = 12, 11.21%) were utilized to perform detection tasks in the fetus head, face, and heart. Additionally, Region Proposal Network (RPN) with bounding-box regression technique was used in (Chen et al., 2020c; Kim et al., 2019a). Faster R-CNN was utilized in (Al-Bander et al., 2020; Kim et al., 2019a; Lin et al., 2019a, 2019b). Further, YOLO (You Only Look Once) models were utilized; YOLOv2 in (Dozen et al., 2020; Fiorentino et al., 2021; Komatsu et al., 2021b) and YOLOv3 in (Pu et al., 2021). Mask R-CNN was utilized only in (Chen et al., 2020b) and Aggregated Residual Visual Block (ARVB) was used only in (Dong et al., 2020). Surprisingly, we have found CNN is used for regression to predict the accurate measurement of HC as seen in (Droste et al., 2020; Fiorentino et al., 2021; Li et al., 2020; Zhang et al., 2020a). We found that, because of the noise and characteristics of US images, DL still relies on traditional feature extraction techniques in image analysis used to optimize the DL frameworks for better performance. For example, Hough transform was utilized with head and abdomen US images as seen in (Deepika et al., 2021; Jang et al., 2017; Kim et al., 2018; Nie et al., 2015a). In addition, image processing techniques such as histogram equalization was used in (Nie et al., 2015a). Besides these tools, Class Activation Maps (CAM) are a helpful classification tool in computer vision that was utilized for brain visualization in (Liu et al., 2020; Xie et al., 2020a), fetal part structure in (Gao and Noble, 2019), and heart chamber view in (Arnaout et al., 2021). DL shows promising results for the identification of certain fetus diseases as seen in the identification of heart diseases (Arnaout et al., 2021; Dinesh Simon and Kavitha, 2021; Dozen et al., 2020; Gong et al., 2020; Komatsu et al., 2021b; Tan et al., 2020; Yang et al., 2020a), brain diseases (Chen et al., 2020b; R. Liu et al., 2020; Sahli et al., 2020; Xie et al., 2020a, 2020b), and growth diseases (Andriani and Mardhiyah, 2019; Selvathi and Chandralekha, 2021; Yekdast, 2019). (Table 9) list of DL studies exhibiting technical novelty that achieved promising results in each category.

Machine learning for fetus health

Most of the research utilizing machine learning for fetal health monitoring was conducted between 2010 and 2015, as seen in (n = 14, 13.8%). Only two studies (Li et al., 2018; Wang et al., 2021) were published in 2018 and 2021, respectively. We found that ML models were used for all fetus organs except for the heart and identification of fetus disease. The most used ML algorithm was Random Forest (RF), as seen in (n = 8, 7.47%) studies. RF was used as a baseline classifier in (Cuingnet et al., 2013; Li et al., 2018; Namburete and Noble, 2013; Yaqub et al., 2013). RF was used with other classifiers as well, with Support Vector Machine

(SVM) in (Ni et al., 2014; Yang et al., 2014) and Probabilistic Boosting Tree (PBT) in (Yaqub et al., 2015). Simple Linear Iterative Clustering (SLIC) was also used in (Rahmatullah et al., 2014). The SVM classifier was used in (n = 6, 5.60%) studies. SVM was used as a baseline classifier in (Lei et al., 2014, 2015; Maraci et al., 2015; Wang et al., 2021) and utilized with other classifiers as seen in (Ni et al., 2014; Yang et al., 2014). AdaBoost was used in (n = 5, 4.67%) studies as a baseline classifier in (Nie et al., 2015b; Rahmatullah et al., 2011a, 2011b, 2012) and to detect the region of interest (ROI) in (Lei et al., 2014). Moreover, ML traditional algorithm still rely on Haar-like features (digital image features used in object recognition) while being utilized with US images as seen in (n = 9, 8.41%) studies (Li et al., 2018; Namburete and Noble, 2013; Ni et al., 2014; Nie et al., 2015b; Rahmatullah et al., 2012; Rahmatullah et al., 2011a, 2011b; Yang et al., 2014; Yaqub et al., 2013). Gaussian Mixture Model (GMM) and Fisher Vector (FV) were used in (n = 3, 2.80%) studies, where GMM is used to simulate the distribution of extracted characteristics throughout the images. Then, FV represents the gradients of the features under the GMM, with respect to the GMM parameters as seen in (Lei et al., 2014; Lin et al., 2019b; Maraci et al., 2015).

DL and ML models were seen in (Sridar et al., 2019), the SVM with decision fusion for classification and features extracted by fine-tuning AlexNet. Finally, we found the first research to utilize ML model in monitoring fetus health was in (Rahmatullah et al., 2011a), which was used to classify abdominal anatomical landmarks.

Artificial Neural Network for fetus health

ANN is the cut edge between ML and DL (Chauhan and Singh, 2019). As seen in Figure 4, ANN was used in (n = 7, 6.5%) studies (Anjit and Rishidas, 2011; Bagi and Shreedhara, 2014; Dave and Nadiad, 2015; Gadagkar and Shreedhara, 2014; Maysanjaya et al., 2014; Rawat et al., 2016; Wee et al., 2010). ANN models were used for the identification of growth diseases (Bagi and Shreedhara, 2014; Gadagkar and Shreedhara, 2014; Rawat et al., 2016), fetus gender (Maysanjaya et al., 2014), facial expressions (Dave and Nadiad, 2015), face anatomical landmarks (Nasal Bone) (Anjit and Rishidas, 2011), and anatomical structures (nuchal translucency (NT)) (Wee et al., 2010). Hence, ANN was not utilized to identify brain or heart structures, nor accompanying disease. Finally, we concluded that the first utilization of ANN to monitor fetus health was in 2010 (Wee et al., 2010); the main goal was to detect NT, which helps early identify Down syndrome. The second utilization of ANN was completed in 2011 (Anjit and Rishidas, 2011) to detect nasal bone. This detection also helps for early identification of Down syndrome.

Reinforcement learning for fetus health

The first attempt to utilize RL to monitor fetus health was in 2019 (Yang et al., 2021a). Dueling Deep Q Networks (DDQN) was employed with RNN-based active termination (AT) to identify brain standard planes (trans-thalamic (TT) and trans-cerebellar (TC)). The second time RL was utilized to monitor fetus health was reported on in a recent study from 2021 (Yang et al., 2021b). A multi-agent RL (MARL) was used in a framework that compromises RNN and includes both neural architecture search (NAS) as well as gradient-based differentiable architecture sampler (GDAS). This framework achieved promising results for identifying the brain standard plane and localize mid-sagittal, transverse (T), coronal (C) planes in volumes, trans-thalamic (TT), trans-ventricular (TV), and trans-cerebellar (TC).

Practical and research implication

The diagnosis of US imaging plays a significant role in clinical patient care. Deep learning, especially the CNN-based model, has lately gained much interest because of its excellent image recognition performance. If CNN were to live up to its potential in radiology, it is expected to aid sonographers and radiologists in achieving diagnostic perfection and improving patient care (Yasaka and Abe, 2018). Diagnostic software based on artificial intelligence is vital in academia and research and has significant media relevance. These systems are largely based on analyzing diagnostic images, such as X-ray, CT, MRI, electrocardiograms, and electroencephalograms. In contrast, US imaging, which is non-invasive, non-expensive, and non-ionizing, has limited AI applications compared to other radiology imaging technologies (ShuoWang et al., 2019).

This survey found one available AI-based application to identify fetal cardiac substructures and indicate structural abnormalities as seen in (Komatsu et al., 2021b). From these findings, many studies achieved promising results in diagnosing fetus diseases or identifying specific fetus landmarks. However, to the best of our knowledge, there is no randomized controlled trial (RCT) or pilot study carried out at a medical center or any adaptation of an AI-based application at any hospitals. This hesitance could be because of the

following challenges (ShuoWang et al., 2019). (1) The present AI system can complete tasks that radiologists are capable of but may make mistakes that a radiologist would not make. For example, radiologists may not see undetectable modifications to the input data. These changes may not be seen by human eyes but would nevertheless impact the result of AI system's categorization. In other words, a tiny variation might cause a deep learning system to reach a different result or judgment. (2) Developers require a specific quantity of reliable and standardized data with an authorized reference standard to train AI systems. Annotated images may become an issue if this is done through retroactive research. Datasets may also be difficult to get as the firms that control them want to keep them private and preserve their intellectual property. Validation of an AI system in the clinic can be difficult as it frequently necessitates multi-institutional collaboration and efficient communication between AI engineers and radiologists. Validating an AI system is also expensive and time-consuming. (3) Finally, when extensive patient databases are involved, ethical and legal concerns may arise.

This survey found some of these challenges were addressed in some studies. For example, applying transfer learning on AI model learned natively on US images and fine-tuning the model on an innovative dataset collected from a different medical center and/or different US equipment. Another way to address the challenges of a small dataset is to employ data augmentation (e.g., tissue deformation, translations, horizontal flips, adding noise, and images enhancement) to boost the generalization capacity of DL models. It is recommended that data augmentation settings should be carefully chosen to replicate ultrasound images changes properly (Akkus et al., 2019). Advanced AI applications are already being used in breast and chest imaging. Large quantities of medical images with strong reference standards are accessible in various fields, allowing the AI system to be trained. Other subspecialties like fetus disease, musculoskeletal disease or interventional radiology are less familiar with utilizing AI. However, it seems that in the future, AI may influence every medical application that utilizes any images (Neri et al., 2019).

This survey found that no standard guideline is being followed or developed that focused on AI in diagnostic imaging to meet clinical setting, evaluation, and requirements. Therefore, the Radiology Editorial Board (Bluemke et al., 2020) has created a list of nine important factors to assist in evaluating AI research as a first step. These recommendations are intended to enhance the validity and usefulness of AI research in diagnostic imaging. These issues are outlined for authors, although manuscript reviewers and readers may find them useful as well: (1) Carefully specify the AI experiment's three image sets (training, validation, and test sets of images). (2) For final statistical reporting, use a separate test set. Overfitting is a problem with AI models, which means they only function for the images they are trained to recognize. It is ideal to utilize an outside collection of images (e.g., the external test set) from another center. (3) Use multivendor images for each step of the AI assessment, if possible (training, validation, test sets). Radiologists understand medical images from one vendor are not identical to theirs; radiomics and AI systems can help in identifying changes in the images. Moreover, multivendor AI algorithms are considerably more interesting than vendor-specific AI algorithms. (4) Always justify the training, validation, and test set sizes. Depending on the application, the number of images needed to train an AI system varies. After just a few hundred images, an AI model may be able to learn image segmentation. (5) Train the AI algorithm using a generally recognized industry standard of reference. For instance, chest radiographs are interpreted by a team of experienced radiologists. (6) Describe any image processing for the AI algorithm. Did the authors manually pick important images, or crop images to a limited field of view? How images are prepared and annotated has a big impact on the radiologist's comprehension of the AI model. (7) Compare AI performance against that of radiology professionals. Competitions and leader boards for the 'best' AI are popular among computer scientists working in the AI field. The area under the receiver operating characteristic curve is often used to compare one AI to another (AUC). On the other hand, physicians are considerably more interested in comparing the AI system to expert readers, when treating a patient. Experienced radiologist readers are recommended to benchmark an algorithm intended to identify radiologic anomalies. (8) Demonstrate the AI algorithm's decision-making process. As previously mentioned, computer scientists working on imaging studies often describe their findings as a single AUC/ACC number. This AUC/ACC is compared to the previous best algorithm, which is a competitor. Unfortunately, the AUC/ACC value has little bearing on clinical medicine on its own. Even with a high AUC/ACC of 0.95, there may be an operating mode in which 99 out of 100 anomalies are overlooked. Many research teams overlay colored probability maps from the AI on the source images to assist doctors in comprehending the AI performance. (9) The AI algorithm should be open source so performance claims may be independently validated; this is already a prevalent recommendation, as this survey found only ($n = 9$, 8.4%) studies (Arnaout et al., 2021; Dozen et al.,

2020; Fiorentino et al., 2021; Gong et al., 2020; Komatsu et al., 2021b; Liu et al., 2021; Skeika et al., 2020; Yang et al., 2019, 2020b) made their works publicly available.

Future work and conclusion

Throughout this survey, various ways that AI techniques have been used to improve fetal care during pregnancy have been discussed, including: (1) determining the existence of a live embryo/fetal and estimate the pregnancy's age, (2) identifying congenital fetal defects and determining the fetus's location, (3) determining the placenta's location, (4) checking for cervix opening or shortening and evaluating fetal development by measuring the quantity of amniotic fluid surrounding the baby, (5) evaluating the fetus's health and seeing whether it is growing properly. Unfortunately, because of both ethical and privacy concerns as well as the reliability of AI decisions, all aspects of these models are not yet utilized as end applications in medical centers.

Undergraduate sonologist education is becoming more essential in medical student careers; consequently, new training techniques are necessary. As seen in other medical fields, AI-based applications and tools were employed for medical and health informatics students (Hasan Sapci and Aylin Sapci, 2020). Moreover, AI-mobile-based applications were useful for medical students, clinicians, and allied health workers alike (Pires et al., 2020). To the best of our knowledge, there are no AI-based mobile applications available to assist radiologist or sonographer students in medical college. Besides that, medical students may face many challenges during fetal US scanning (Bahner et al., 2012). Therefore, our future work will propose an AI-mobile-based application to help radiology or sonography students identify features in fetus US image and answer the following research questions: (1) How is the lightweight model efficient and accurate when localizing the fetal head, abdomen, femur, and thorax from the first to the third trimester; (2) How efficiently and accurately will the lightweight model identify the fetal gestational age (GA) from the first to the third trimester?; (3) How will inter-rater reliability tests validate the proposed model and compare the obtained result with experts and students using the intra-class correlation coefficient (ICC) (Bartko, 1966). We have concluded that AI techniques utilized US images to monitor fetal health from different aspects through various GA. Out of 107 studies, DL was the most widely used model, followed by ML, ANN, and RL. These models were used to implement various tasks, including classification, segmentation, object detection, and regression. We found that even the most recent studies rely on the 2D US followed by 3D, and that 4D is rarely used. Furthermore, we found that most of the work targeted the fetus head followed by the body, heart, face, and abdomen. We identified the lead institutes in this field and their research. This survey discusses the availability of the dataset for each category and highlights their promising results. In addition, we analyzed each study independently and provided observations for the future reader. All optimization and feature extraction methods were reported for each study to highlight the unique contribution. Moreover, DL novel and unique works were highlighted for each category. The research and practical implications were discussed, and prospective research and clinical practitioner recommendations were provided. Finally, a future research direction has been proposed to answer the gap in this survey.

Limitations of the study

To the best of our knowledge, this survey is the first to explore all AI techniques implemented to provide fetus health care and monitoring during different pregnancy trimesters. This survey may be considered comprehensive as it does not focus on specific AI branches or diseases. Rather, it provides a holistic view of AI's role in monitoring fetuses via ultrasound images. This survey will benefit both readers, first medical professionals to have sight about the current AI technology in a development stage and assist data scientists in overcoming existing research and implication problems. Secondly, for data scientists to continue future investigating in developing lightweight models and overcoming problems that slow the transformation of this technology based on a solid guideline proposed by the medical professionals. It may be considered a robust and high-quality survey. Because the most prominent health and information technology databases were searched using a well-developed search query, the search was sensitive and exact. Lastly, the utilization of techniques like scanning gray literature databases (Google Scholar), biomedical literature database (PubMed, Embase) World's leading citation databases (Web of Science), Subject Specific Databases (PsycINFO), and world's leading source for scientific, technical, and medical research (ScienceDirect, IEEE explore, ACM Library) indicates that this study's risk of publication bias is low. A comprehensive systematic survey was conducted. However, the search within the databases was conducted between 22nd and 23rd June 202, so we might miss some new studies. Besides that, the words "pregnancy" or "pregnant" or

“uterus” were not used as search terms to identify relevant papers. Therefore, some fetus health monitoring studies might have been missed, lowering the overall number of studies. In addition, this survey focused more on the clinical rather than the technical; therefore, some technical details may have been missed. Lastly, only English studies were included in the search. As a result, research written in other languages was excluded.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
 - Lead contact
 - Materials availability
 - Data and code availability
- [METHODS DETAILS](#)
 - Eligibility criteria
 - Search strategy
 - Evaluating tool and reporting standard
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104713>.

ACKNOWLEDGMENTS

Open Access funding provided by the Qatar National Library.

AUTHOR CONTRIBUTIONS

M.A., M.H., and A.A.A conducted the protocol and drafting of the systematic review. M.A., and U.S. performed literature search and data retrieval. Marco Agus and K.A.A., provided guidance and made suggestions on machine learning algorithms. M.A., and Mohammed Anbar interpreted the analysis result. M.H., K.A., and M.M., contributed to the subsequent revisions of the manuscripts. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 30, 2022

Revised: June 9, 2022

Accepted: June 28, 2022

Published: August 19, 2022

REFERENCES

- Abramowicz, J. (2015). ALARA: the clinical view. *Ultrasound Med. Biol.* 41, S102. <https://doi.org/10.1016/j.ultrasmedbio.2014.12.677>.
- Abramowicz, J.S. (2010). Nonmedical use of ultrasound: bioeffects and safety risk. *Ultrasound Med. Biol.* 36, 1213–1220. <https://doi.org/10.1016/j.ultrasmedbio.2010.04.003>.
- Aji, C.P., Fatoni, M.H., and Sardjono, T.A. (2019). Automatic measurement of fetal head circumference from 2-dimensional ultrasound. In 2019 International Conference On Computer Engineering, Network, and Intelligent Multimedia, CENIM 2019 - Proceeding, 2019-Novem, 1–5. <https://doi.org/10.1109/CENIM48368.2019.8973258>.
- Akkus, Z., Cai, J., Boonrod, A., Zeinoddini, A., Weston, A.D., Philbrick, K.A., and Erickson, B.J. (2019). A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow. *J. Am. Coll. Radiol.* 16, 1318–1328. <https://doi.org/10.1016/j.jacr.2019.06.004>.
- Al-Bander, B., Alzahrani, T., Alzahrani, S., Williams, B.M., and Zheng, Y. (2020). Improving fetal head contour detection by object localisation with deep learning. *Commun. Comput. Inf. Sci.* 1065, 142–150. https://doi.org/10.1007/978-3-030-39343-4_12.
- Al-yousif, S., Jaenul, A., Al-Dayyeni, W., Alamoodi, A., Jabori, I., Tahir, N.M., Alrawi, A.A.A., Cömert, Z., Al-shareefi, N.A., and Saleh, A.H. (2021). A systematic review of automated preprocessing, feature extraction and classification of cardiocography. *PeerJ Comput. Sci.* 7, 1–37. <https://doi.org/10.7717/peerj-cs.452>.
- Alzubaidi, M., Zubaydi, H.D., Bin-Salem, A.A., Abd-Alrazaq, A.A., Ahmed, A., and Househ, M. (2021b). Role of deep learning in early detection of COVID-19: scoping review. *Comput. Methods Programs Biomed.* 1, 100025. <https://doi.org/10.1016/j.cmpbup.2021.100025>.
- Alzubaidi, M.S., Shah, U., Dhia Zubaydi, H., Dolaat, K., Abd-Alrazaq, A.A., Ahmed, A., and Househ, M. (2021a). The role of neural network for the detection of Parkinson’s disease: a scoping

- review. *Healthcare (Switzerland)* 9. <https://doi.org/10.3390/healthcare9060740>.
- Andriani, F., and Mardiyah, I. (2019). Blighted Ovum detection using convolutional neural network. In *AIP Conference Proceedings, 2084*. <https://doi.org/10.1063/1.5094276>.
- Anjit, T.A., and Rishidas, S. (2011). Identification of nasal bone for the early detection of down syndrome using Back Propagation Neural Network. In *ICCCSP 2011 - 2011 International Conference on Communications and Signal Processing*, pp. 136–140. <https://doi.org/10.1109/ICCCSP.2011.5739286>.
- Arjunan, S.P., and Thomas, M.C. (2020). A review of ultrasound imaging techniques for the detection of down syndrome. *Irbm* 41, 115–123. <https://doi.org/10.1016/j.irbm.2019.10.004>.
- Arnaout, R., Curran, L., Zhao, Y., Levine, J.C., Chinn, E., and Moon-Grady, A.J. (2021). An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. *Nat. Med.* 27, 882–891. <https://doi.org/10.1038/s41591-021-01342-5>.
- Asgari Taghanaki, S., Abhishek, K., Cohen, J.P., Cohen-Adad, J., and Hamarneh, G. (2021). Deep semantic segmentation of natural and medical images: a review. *Artif. Intell. Rev.* 54, 137–178. <https://doi.org/10.1007/s10462-020-09854-1>.
- Avola, D., Cinque, L., Fagioli, A., Foresti, G., and Mecca, A. (2021). Ultrasound medical imaging techniques. *ACM Comput. Surv.* 54. <https://doi.org/10.1145/3447243>.
- Bagi, K.S., and Shreedhara, K.S. (2014). Biometric measurement and classification of IUGR using neural networks. In *Proceedings of 2014 International Conference on Contemporary Computing and Informatics, IC3I*, pp. 157–161. <https://doi.org/10.1109/IC3I.2014.7019613>.
- Bahner, D.P., Jasne, A., Boore, S., Mueller, A., and Cortez, E. (2012). The ultrasound challenge: a novel approach to medical student ultrasound education. *J. Ultrasound Med.* 31, 2013–2016. <https://doi.org/10.7863/jum.2012.31.12.2013>.
- Balayla, J., and Shrem, G. (2019). Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis. *Arch. Gynecol. Obstet.* 300, 7–14. <https://doi.org/10.1007/s00404-019-05151-7>.
- Bali, A., and Singh, S.N. (2015). A review on the strategies and techniques of image segmentation. In *International Conference on Advanced Computing and Communication Technologies, ACCT, 2015-April*, pp. 113–120. <https://doi.org/10.1109/ACCT.2015.63>.
- Bartko, J.J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychol. Rep.* 19, 3–11. <https://doi.org/10.2466/pr0.1966.19.1.3>.
- Bethune, M., Alibrahim, E., Davies, B., and Yong, E. (2013). A pictorial guide for the second trimester ultrasound. *Australas. J. Ultrasound Med.* 16, 98–113. <https://doi.org/10.1002/j.2205-0140.2013.tb00106.x>.
- Bin-Salem, A.A., Zubaydi, H.D., Alzubaidi, M., Ul Abideen Tariq, Z., and Naeem, H. (2022). A scoping review on COVID-19's early detection using deep learning model and computed tomography and ultrasound. *Trait. Du. Signal* 39, 205–219. <https://doi.org/10.18280/ts.390121>.
- Bluemke, D.A., Moy, L., Bredella, M.A., Ertl-Wagner, B.B., Fowler, K.J., Goh, V.J., Halpern, E.F., Hess, C.P., Schiebler, M.L., and Weiss, C.R. (2020). Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers-from the Radiology Editorial Board. *Radiology* 294, 487–489. <https://doi.org/10.1148/radiol.2019192515>.
- Brahma, K., Kumar, V., Samir, A.E., Chandrakasan, A.P., and Eldar, Y.C. (2021). Efficient binary cnn for medical image segmentation. In *Proceedings - International Symposium on Biomedical Imaging, 2021-April*, pp. 817–821. <https://doi.org/10.1109/ISBI48211.2021.9433901>.
- Budd, S., Sinclair, M., Khanal, B., Matthew, J., Lloyd, D., Gomez, A., Toussaint, N., Robinson, E.C., and Kainz, B. (2019). Confident head circumference measurement from ultrasound with real-time feedback for sonographers. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11767 LNCS, pp. 683–691. https://doi.org/10.1007/978-3-030-32251-9_75.
- Burgos-Artizzu, X.P., Coronado-Gutiérrez, D., Valenzuela-Alcaraz, B., Bonet-Carne, E., Eixarch, E., Crispí, F., and Gratacós, E. (2020). Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Sci. Rep.* 10, 1–12. <https://doi.org/10.1038/s41598-020-67076-5>.
- Cai, Y., Droste, R., Sharma, H., Chatelain, P., Drukker, L., Papageorghiou, A.T., and Noble, J.A. (2020). Spatio-temporal visual attention modelling of standard biometry plane-finding navigation. *Med. Image Anal.* 65, 101762. <https://doi.org/10.1016/j.media.2020.101762>.
- Cai, Y., Sharma, H., Chatelain, P., and Noble, J.A. (2018). Multi-task SonoEyeNet: detection of fetal standardized planes assisted by generated sonographer attention maps. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11070 LNCS, pp. 871–879. https://doi.org/10.1007/978-3-030-00928-1_98.
- Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., Gallivanone, F., Cozzi, A., D'Amico, N.C., and Sardaneli, F. (2021). AI applications to medical images: from machine learning to deep learning. *Phys. Med.* 83, 9–24. <https://doi.org/10.1016/j.ejmp.2021.02.006>.
- Cerrolaza, J.J., Sinclair, M., Li, Y., Gomez, A., Ferrante, E., Matthew, J., Gupta, C., Knight, C.L., and Rueckert, D. (2018). Deep learning with ultrasound physics for fetal skull segmentation. In *Proceedings - International Symposium on Biomedical Imaging, 2018-April*, pp. 564–567. <https://doi.org/10.1109/ISBI.2018.8363639>.
- Chauhan, N.K., and Singh, K. (2019). A review on conventional machine learning vs deep learning. In *2018 International Conference on Computing, Power and Communication Technologies, GUCON 2018*, pp. 347–352. <https://doi.org/10.1109/GUCON.2018.8675097>.
- Cheikh, G.A., Mbacke, A.B., and Ndiaye, S. (2020). Deep learning in medical imaging survey. *CEUR Workshop Proc.* 2647, 111–127. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6945006/>.
- Chen, C., Yang, X., Huang, R., Shi, W., Liu, S., Lin, M., Huang, Y., Yang, Y., Zhang, Y., Luo, H., et al. (2020c). Region proposal network with Graph prior and l₀-balance loss for landmark detection in 3D ultrasound. In *Proceedings - International Symposium on Biomedical Imaging, 2020-April*, pp. 1829–1833. <https://doi.org/10.1109/ISBI45749.2020.9098368>.
- Chen, H., Dou, Q., Ni, D., Cheng, J.Z., Qin, J., Li, S., and Heng, P.A. (2015). Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. *Lect. Notes Comput. Sci.* 9349, 507–514. https://doi.org/10.1007/978-3-319-24553-9_62.
- Chen, H., Ni, D., Yang, X., Li, S., and Heng, P.A. (2014). Fetal abdominal standard plane localization through representation learning with knowledge transfer. *Lect. Notes Comput. Sci.* 8679, 125–132. https://doi.org/10.1007/978-3-319-10581-9_16.
- Chen, H., Wu, L., Dou, Q., Qin, J., Li, S., Cheng, J.Z., Ni, D., and Heng, P.A. (2017). Ultrasound standard plane detection using a composite neural network framework. *IEEE Trans. Cybern.* 47, 1576–1583. <https://doi.org/10.1109/TCYB.2017.2685080>.
- Chen, P., Chen, Y., Deng, Y., Wang, Y., He, P., Lv, X., and Yu, J. (2020a). A preliminary study to quantitatively evaluate the development of maturation degree for fetal lung based on transfer learning deep model from ultrasound images. *Int. J. Comput. Assist. Radiol. Surg.* 15, 1407–1415. <https://doi.org/10.1007/s11548-020-02211-1>.
- Chen, X., He, M., Dan, T., Wang, N., Lin, M., Zhang, L., Xian, J., Cai, H., and Xie, H. (2020b). Automatic measurements of fetal lateral ventricles in 2D ultrasound images using deep learning. *Front. Neurol.* 11, 526. <https://doi.org/10.3389/fneur.2020.00526>.
- Chen, Z., Liu, Z., Du, M., and Wang, Z. (2021a). Artificial intelligence in obstetric ultrasound: an update and future applications. *Front. Med.* 8, 1431. <https://doi.org/10.3389/fmed.2021.733468>.
- Chen, Z., Wang, Z., Du, M., and Liu, Z. (2021b). Artificial intelligence in the assessment of female reproductive function using ultrasound: a review. *J. Ultrasound Med.* <https://doi.org/10.1002/jum.15827>.
- Correa, P.J., Palmeiro, Y., Soto, M.J., Ugarte, C., and Illanes, S.E. (2016). Etiopathogenesis, prediction, and prevention of preeclampsia. *Hypertens. Pregnancy* 35, 280–294. <https://doi.org/10.1080/10641955.2016.1181180>.
- Cuingnet, R., Somphone, O., Mory, B., Prevost, R., Yaqub, M., Napolitano, R., Papageorghiou, A., Roundhill, D., Noble, J.A., and Ardon, R. (2013). Where is my baby? A fast fetal head auto-alignment in 3D-ultrasound. In *Proceedings - International Symposium on Biomedical Imaging*, pp. 768–771. <https://doi.org/10.1109/ISBI.2013.6556588>.
- Dave, P.R., and Nadiad, M.S.B. (2015). Facial expressions extraction from 3D sonography

- images. In Proceedings of 2015 IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2015, pp. 1–6. <https://doi.org/10.1109/ICECCT.2015.7226070>.
- Davidson, L., and Boland, M.R. (2021). Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes. *Brief. Bioinformatics*. <https://doi.org/10.1093/bib/bbaa369>.
- Deepika, P., Suresh, R.M., and Pabitha, P. (2021). Defending against Child Death: deep learning-based diagnosis method for abnormal identification of fetus ultrasound Images. *Comput. Intell.* 37, 128–154. <https://doi.org/10.1111/coin.12394>.
- Deo, R.C. (2015). Machine learning in medicine. *Circulation* 132, 1920–1930. <https://doi.org/10.1161/CIRCULATIONAHA.115.001593>.
- Desai, A., Chauhan, R., and Sivaswamy, J. (2020). Image segmentation using hybrid representations. In Proceedings - International Symposium on Biomedical Imaging, 2020-April, pp. 1513–1516. <https://doi.org/10.1109/ISBI45749.2020.9098463>.
- Dinesh Simon, M., and Kavitha, A.R. (2021). Ultrasonic detection of down syndrome using Multiscale Quantiser with convolutional neural network. In Computational Optimization Techniques and Applications (IntechOpen). <https://doi.org/10.5772/intechopen.96502>.
- Dong, J., Liu, S., Liao, Y., Wen, H., Lei, B., Li, S., and Wang, T. (2020). A generic quality control framework for fetal ultrasound cardiac four-chamber planes. *IEEE Journal of Biomedical and Health Informatics* 24, 931–942. <https://doi.org/10.1109/JBHI.2019.2948316>.
- Dowdy, D. (2016). Keepsake ultrasound: taking another look. *J. Radiol. Nurs.* 35, 119–132. <https://doi.org/10.1016/j.jradnu.2016.02.006>.
- Dozen, A., Komatsu, M., Sakai, A., Komatsu, R., Shozu, K., Machino, H., Yasutomi, S., Arakaki, T., Asada, K., Kaneko, S., et al. (2020). Image segmentation of the ventricular septum in fetal cardiac ultrasound videos based on deep learning using time-series information. *Biomolecules* 10, 1–17. <https://doi.org/10.3390/biom10111526>.
- Driscoll, D.A., and Gross, S. (2009). Prenatal screening for Aneuploidy. *N. Engl. J. Med.* 360, 2556–2562. <https://doi.org/10.1056/nejmcp0900134>.
- Droste, R., Chatelain, P., Drukker, L., Sharma, H., Papageorgiou, A.T., and Noble, J.A. (2020). Discovering salient anatomical landmarks by predicting human Gaze. In Proceedings - International Symposium on Biomedical Imaging, 2020-April, pp. 1711–1714. <https://doi.org/10.1109/ISBI45749.2020.9098505>.
- Fathimathu Joharah, S., and Mohideen, K. (2020). Automatic detection of fetal ultrasound image using multi-task deep learning. *J. Crit. Rev.* 7, 987–993. <https://doi.org/10.31838/jcr.07.07.181>.
- Fatima, M., and Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *J. Intell. Learn. Syst. Appl.* 09, 1–16. <https://doi.org/10.4236/jilsa.2017.91001>.
- Fiorentino, M.C., Moccia, S., Capparuccini, M., Giamberini, S., and Frontoni, E. (2021). A regression framework to head-circumference delineation from US fetal images. *Comput. Methods Progr. Biomed.* 198, 105771. <https://doi.org/10.1016/j.cmpb.2020.105771>.
- Fujita, H. (2020). AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. *Radiol. Phys. Technol.* 13, 6–19. <https://doi.org/10.1007/s12194-019-00552-4>.
- Gadagkar, A.V., and Shreedhara, K.S. (2014). Features based IUGR diagnosis using variational level set method and classification using artificial neural networks. In Proceedings - 2014 5th International Conference on Signal and Image Processing, ICSIP 2014, pp. 303–309. <https://doi.org/10.1109/ICSIP.2014.54>.
- Gao, Y., and Noble, J.A. (2019). Learning and understanding deep spatio-temporal representations from free-hand fetal ultrasound sweeps. *Lect. Notes Comput. Sci.* 11768, 299–308. https://doi.org/10.1007/978-3-030-32254-0_34.
- Garcia-Canadilla, P., Sanchez-Martinez, S., Crispi, F., and Bijmens, B. (2020). Machine learning in fetal cardiology: what to expect. *Fetal Diagn. Ther.* 47, 363–372. <https://doi.org/10.1159/000505021>.
- Gong, Y., Zhang, Y., Zhu, H., Lv, J., Cheng, Q., Zhang, H., He, Y., and Wang, S. (2020). Fetal congenital heart disease echocardiogram screening based on dqacnn: adversarial one-class classification combined with video transfer learning. *IEEE Trans. Med. Imag.* 39, 1206–1222. <https://doi.org/10.1109/TMI.2019.2946059>.
- Handelman, G.S., Kok, H.K., Chandra, R.V., Razavi, A.H., Lee, M.J., and Asadi, H. (2018). eDoctor: machine learning and the future of medicine. *J. Intern. Med.* 284, 603–619. <https://doi.org/10.1111/joim.12822>.
- Hartkopf, J., Moser, J., Schleger, F., Preissl, H., and Keune, J. (2019). Changes in event-related brain responses and habituation during child development – a systematic literature review. *Clin. Neurophysiol.* 130, 2238–2254. <https://doi.org/10.1016/j.clinph.2019.08.029>.
- Hasan Sapci, A., and Aylin Sapci, H. (2020). Artificial intelligence education and tools for medical and health informatics students: systematic review. *JMIR Med. Educ.* 6. <https://doi.org/10.2196/19285>.
- Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron* 95, 245–258. <https://doi.org/10.1016/j.neuron.2017.06.011>.
- Hesamian, M.H., Jia, W., He, X., and Kennedy, P. (2019). Deep learning techniques for medical image segmentation: achievements and challenges. *J. Digit. Imag.* 32, 582–596. <https://doi.org/10.1007/s10278-019-00227-x>.
- Huang, Q., and Zeng, Z. (2017). A review on real-time 3D ultrasound imaging technology. *BioMed Res. Int.* <https://doi.org/10.1155/2017/6027029>.
- Huang, R., Xie, W., and Alison Noble, J. (2018). VP-Nets: efficient automatic localization of key brain structures in 3D fetal neurosonography. *Med. Image Anal.* 47, 127–139. <https://doi.org/10.1016/j.media.2018.04.004>.
- Iftikhar, P.M., Kuijpers, M.V., Khayyat, A., Iftikhar, A., and DeGouvoua De Sa, M. (2020). Artificial intelligence: a new paradigm in obstetrics and gynecology research and clinical practice. *Cureus*. <https://doi.org/10.7759/cureus.7124>.
- Jang, J., Park, Y., Kim, B., Lee, S.M., Kwon, J.Y., and Seo, J.K. (2017). Automatic estimation of fetal abdominal circumference from ultrasound images. *IEEE J. Biomed. Health Inform.* 22, 1512–1520. <https://doi.org/10.1109/JBHI.2017.2776116>.
- Kaur, P., Singh, G., and Kaur, P. (2017). A review of denoising medical images using machine learning approaches. *Curr. Med. Imag. Rev.* 14, 675–685. <https://doi.org/10.15173405613666170428154156>.
- Kim, B., Kim, K.C., Park, Y., Kwon, J.Y., Jang, J., and Seo, J.K. (2018). Machine-learning-based automatic identification of fetal abdominal circumference from ultrasound images. *Physiol. Meas.* 39, 105007. <https://doi.org/10.1088/1361-6579/aae255>.
- Kim, H.P., Lee, S.M., Kwon, J.Y., Park, Y., Kim, K.C., and Seo, J.K. (2019a). Automatic evaluation of fetal head biometry from ultrasound images using machine learning. *Physiol. Meas.* 40, 65009. <https://doi.org/10.1088/1361-6579/ab21ac>.
- Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H.J., and Kim, N. (2019b). Deep learning in medical imaging. *Neurospine* 16, 657–668. <https://doi.org/10.14245/ns.1938396.198>.
- Kokol, P., Blažun Vošner, H., and Završnik, J. (2021). Application of bibliometrics in medicine: a historical bibliometrics analysis. *Health Inf. Libr. J.* 38, 125–138. <https://doi.org/10.1111/hir.12295>.
- Komatsu, M., Sakai, A., Dozen, A., Shozu, K., Yasutomi, S., Machino, H., Asada, K., Kaneko, S., and Hamamoto, R. (2021a). Towards clinical application of artificial intelligence in ultrasound imaging. *Biomedicine* 9. <https://doi.org/10.3390/biomedicine9070720>.
- Komatsu, M., Sakai, A., Komatsu, R., Matsuoka, R., Yasutomi, S., Shozu, K., Dozen, A., Machino, H., Hidaka, H., Arakaki, T., et al. (2021b). Detection of cardiac structural abnormalities in fetal ultrasound videos using deep learning. *Appl. Sci.* 11, 1–12. <https://doi.org/10.3390/app11010371>.
- Kurjak, A., Miskovic, B., Andonotopo, W., Stanojevic, M., Azumendi, G., and Vrcic, H. (2007). How useful is 3D and 4D ultrasound in perinatal medicine? *J. Perinat. Med.* 35, 10–27. <https://doi.org/10.1515/JPM.2007.002>.
- Larsen, E.C., Christiansen, O.B., Kolte, A.M., and Macklon, N. (2013). New insights into mechanisms behind miscarriage. *BMC Med.* 11. <https://doi.org/10.1186/1741-7015-11-154>.
- Lei, B., Tan, E.L., Chen, S., Zhuo, L., Li, S., Ni, D., and Wang, T. (2015). Automatic recognition of fetal facial standard plane in ultrasound image via Fisher vector. *PLoS One* 10, e0121838. <https://doi.org/10.1371/journal.pone.0121838>.

- Lei, B., Zhuo, L., Chen, S., Li, S., Ni, D., and Wang, T. (2014). Automatic recognition of fetal standard plane in ultrasound image. In IEEE 11th International Symposium on Biomedical Imaging (ISBI), pp. 85–88. <https://doi.org/10.1109/isbi.2014.6867815>.
- Li, J., Wang, Y., Lei, B., Cheng, J.Z., Qin, J., Wang, T., Li, S., and Ni, D. (2018). Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting. *IEEE J. Biomed. Health Inform.* 22, 215–223. <https://doi.org/10.1109/JBHI.2017.2703890>.
- Li, P., Zhao, H., Liu, P., and Cao, F. (2020). Automated measurement network for accurate segmentation and parameter modification in fetal head ultrasound images. *Med. Biol. Eng. Comput.* 58, 2879–2892. <https://doi.org/10.1007/s11517-020-02242-5>.
- Li, Y., Xu, R., Ohya, J., and Iwata, H. (2017). Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 1485–1488. <https://doi.org/10.1109/EMBC.2017.8037116>.
- Lin, Z., Lei, B., Jiang, F., Ni, D., Chen, S., Li, S., and Wang, T. (2019a). Quality assessment of fetal head ultrasound images based on faster R-CNN. *Chin. J. Biomed. Eng.* 38, 392–400. <https://doi.org/10.3969/j.issn.0258-8021.2019.04.002>.
- Lin, Z., Li, S., Ni, D., Liao, Y., Wen, H., Du, J., Chen, S., Wang, T., and Lei, B. (2019b). Multi-task learning for quality assessment of fetal head ultrasound images. *Med. Image Anal.* 58, 101548. <https://doi.org/10.1016/j.media.2019.101548>.
- Liu, R., Liua, M., Sheng, B., Li, H., Li, P., Song, H., Zhang, P., Jiang, L., and Shen, D. (2021). NHBS-net: a feature fusion attention network for ultrasound neonatal hip bone segmentation. *IEEE Trans. Med. Imaging.* <https://doi.org/10.1109/TMI.2021.3087857>.
- Liu, R., Zou, B., Zhang, H., Ye, J., Han, C., Zhan, N., Yang, Y., Zhang, H., Chen, F., Hua, S., and Guo, J. (2020). Automated fetal lateral ventricular width estimation from prenatal ultrasound based on deep learning algorithms. Preprints. Authorea. <https://doi.org/10.22541/AU.158872369.94920854>.
- Liu, T., Xu, M., Zhang, Z., Dai, C., Wang, H., Zhang, R., Shi, L., and Wu, S. (2019). Direct detection and measurement of nuchal translucency with neural networks from ultrasound images. *Lect. Notes Comput. Sci.* 11798, 20–28. https://doi.org/10.1007/978-3-030-32875-7_3.
- Looney, P., Yin, Y., Collins, S.L., Nicolaides, K.H., Plasencia, W., Molloholli, M., Natsis, S., and Stevenson, G.N. (2021). Fully automated 3-D ultrasound segmentation of the placenta, amniotic fluid, and fetus for early pregnancy assessment. *IEEE Trans. Ultrason. Ferroelectrics Freq. Control* 68, 2038–2047. <https://doi.org/10.1109/TUFFC.2021.3052143>.
- Loweckamp, B.C., Chen, D.T., Ibáñez, L., and Blezek, D. (2013). The design of simpleITK. *Front. Neuroinf.* 7, 45. <https://doi.org/10.3389/fninf.2013.00045>.
- Mack, L. (2017). What are the differences between 2D, 3D and 4D ultrasounds scans?. <https://www.forlongmedical.com/What-are-the-differences-between-2D-3D-and-4D-ultrasounds-scans-id529841.html>.
- Maraci, M.A., Napolitano, R., Papageorghiou, A., and Noble, J.A. (2015). Fisher vector encoding for detecting objects of interest in ultrasound videos. In Proceedings - International Symposium on Biomedical Imaging, 2015-July, pp. 651–654. <https://doi.org/10.1109/ISBI.2015.7163957>.
- Maraci, M.A., Yaqub, M., Craik, R., Beriwal, S., Self, A., von Dadelszen, P., Papageorghiou, A., and Noble, J.A. (2020). Toward point-of-care ultrasound estimation of fetal gestational age from the trans-cerebellar diameter using CNN-based ultrasound image analysis. *J. Med. Imaging* 7, 1. <https://doi.org/10.1117/1.jmi.7.1.014501>.
- Masselli, G., Brunelli, R., Monti, R., Guida, M., Laghi, F., Casciani, E., Poletti, E., and Gualdi, G. (2014). Imaging for acute pelvic pain in pregnancy. *Insights Imaging* 5, 165–181. <https://doi.org/10.1007/s13244-014-0314-8>.
- Maysanjaya, I.M.D., Nugroho, H.A., and Setiawan, N.A. (2014). The classification of fetus gender on ultrasound images using learning vector quantization (LVQ). In Proceeding - 2014 Makassar International Conference on Electrical Engineering and Informatics, MICEEI, pp. 150–155. <https://doi.org/10.1109/MICEEI.2014.7067329>.
- Miller, D.D., and Brown, E.W. (2018). Artificial intelligence in medical practice: the question to the answer? *Am. J. Med.* 131, 129–133. <https://doi.org/10.1016/j.amjmed.2017.10.035>.
- Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J.T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings Bioinf.* 19, 1236–1246. <https://doi.org/10.1093/bib/bbx044>.
- Miyagi, Y., Hata, T., Bouno, S., Koyanagi, A., and Miyake, T. (2021). Recognition of facial expression of fetuses by artificial intelligence (AI). *J. Perinat. Med.* 49, 596–603. <https://doi.org/10.1515/jpm-2020-0537>.
- Naeem, H., and Bin-Salem, A.A. (2021). A CNN-LSTM network with multi-level feature extraction-based approach for automated detection of coronavirus from CT scan and X-ray images. *Appl. Soft Comput.* 113. <https://doi.org/10.1016/j.asoc.2021.107918>.
- Namburete, A.I.L., and Noble, J.A. (2013). Fetal cranial segmentation in 2D ultrasound images using shape properties of pixel clusters. In Proceedings - International Symposium on Biomedical Imaging, pp. 720–723. <https://doi.org/10.1109/ISBI.2013.6556576>.
- Namburete, A.I.L., Xie, W., Yaqub, M., Zisserman, A., and Noble, J.A. (2018). Fully-automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning. *Med. Image Anal.* 46, 1–14. <https://doi.org/10.1016/j.media.2018.02.006>.
- Neri, E., de Souza, N., Brady, A., Bayarri, A.A., Becker, C.D., Coppola, F., and Visser, J. (2019). What the radiologist should know about artificial intelligence – an ESR white paper. *Insights Imaging* 10. <https://doi.org/10.1186/s13244-019-0738-2>.
- Ni, D., Yang, X., Chen, X., Chin, C.T., Chen, S., Heng, P.A., Li, S., Qin, J., and Wang, T. (2014). Standard plane localization in ultrasound by radial component model and selective search. *Ultrasound Med. Biol.* 40, 2728–2742. <https://doi.org/10.1016/j.ultrasmedbio.2014.06.006>.
- Nie, S., Yu, J., Chen, P., Zhang, J., and Wang, Y. (2015a). A Novel Method with a Deep Network and Directional Edges for Automatic Detection of a Fetal Head. In 2015 23rd European Signal Processing Conference (EUSIPCO), pp. 654–658. <https://doi.org/10.1109/EUSIPCO.2015.7362464>.
- Nie, S., Yu, J., Wang, Y., Zhang, J., and Chen, P. (2015b). Shape model and marginal space of 3D ultrasound volume data for automatically detecting a fetal head. In ICALIP 2014 - 2014 International Conference on Audio, Language and Image Processing, pp. 681–685. <https://doi.org/10.1109/ICALIP.2014.7009881>.
- Patra, A., Huang, W., and Noble, J.A. (2017). Learning spatio-temporal aggregation for fetal heart analysis in ultrasound video. *Lect. Notes Comput. Sci.* 10553, 276–284. https://doi.org/10.1007/978-3-319-67558-9_32.
- Payan, C., Abraham, O., and Garnier, V. (2018). Ultrasonic methods. In Non-destructive Testing and Evaluation of Civil Engineering Structures (Elsevier), pp. 21–85. <https://doi.org/10.1016/B978-1-78548-229-8.50002-9>.
- Perez-Gonzalez, J., Hevia Montiel, N., and Bañuelos, V.M. (2020). Deep learning spatial compounding from multiple fetal head ultrasound acquisitions. *Lect. Notes Comput. Sci.* 12437, 305–314. https://doi.org/10.1007/978-3-030-60334-2_30.
- Pieper, S., Halle, M., and Kikinis, R. (2004). 3D Slicer. In 2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano, 12004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano, pp. 632–635. <https://doi.org/10.1007/s00113-019-0654-4>.
- Pires, I.M., Marques, G., Garcia, N.M., Flórez-revuelta, F., Ponciano, V., and Oniani, S. (2020). A research on the classification and applicability of the mobile health applications. *J. Personal. Med.* 10. <https://doi.org/10.3390/jpm10010011>.
- Prieto, J.C., Shah, H., Rosenbaum, A., Jiang, X., Musonda, P., Price, J., Stringer, E.M., Vwalika, B., Stamilio, D., and Stringer, J. (2021). An automated framework for image classification and segmentation of fetal ultrasound images for gestational age estimation. *Proc. SPIE-Int. Soc. Opt. Eng.* 11596, 55. <https://doi.org/10.1117/12.2582243>.
- Pu, B., Zhu, N., Li, K., and Li, S. (2021). Fetal cardiac cycle detection in multi-resource echocardiograms using hybrid classification framework. *Future Generat. Comput. Syst.* 115, 825–836. <https://doi.org/10.1016/j.future.2020.09.014>.
- Qiao, D., and Zulkernine, F. (2020). Dilated squeeze-and-excitation U-net for fetal ultrasound image segmentation. In IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB, pp. 1–7. <https://doi.org/10.1109/CIBCB48159.2020.9277667>.

- Qu, R., Xu, G., Ding, C., Jia, W., and Sun, M. (2020a). Deep learning-based methodology for recognition of fetal brain standard scan planes in 2D ultrasound images. *IEEE Access* 8, 44443–44451. <https://doi.org/10.1109/ACCESS.2019.2950387>.
- Qu, R., Xu, G., Ding, C., Jia, W., and Sun, M. (2020b). Standard plane identification in fetal brain ultrasound scans using a differential convolutional neural network. *IEEE Access* 8, 83821–83830. <https://doi.org/10.1109/ACCESS.2020.2991845>.
- Raef, B., and Ferdousi, R. (2019). A review of machine learning approaches in assisted reproductive technologies. *Acta Inf. Med.* 27, 205–211. <https://doi.org/10.5455/aim.2019.27.205-211>.
- Rahmatnllah, B., Papageorghiou, A.T., and Alison Noble, J. (2012). Image analysis using machine learning: anatomical landmarks detection in fetal ultrasound images. *CompSAC*, 354–355. <https://doi.org/10.1109/COMPSAC.2012.52>.
- Rahmatullah, B., Papageorghiou, A., and Noble, J.A. (2011a). Automated selection of standardized planes from ultrasound volume. *Lect. Notes Comput. Sci.* 7009, 35–42. https://doi.org/10.1007/978-3-642-24319-6_5.
- Rahmatullah, B., Sarris, I., Papageorghiou, A., and Noble, J.A. (2011b). Quality control of fetal ultrasound images: detection of abdomen anatomical landmarks using AdaBoost. In *Proceedings - International Symposium on Biomedical Imaging*, pp. 6–9. <https://doi.org/10.1109/ISBI.2011.5872342>.
- Rahmatullah, R., Ma'Sum, M.A., Aprinaldi, Mursanto, P., and Wiwiko, B. (2014). Automatic fetal organs segmentation using multilayer super pixel and image moment feature. In *Proceedings - ICACIS 2014: 2014 International Conference on Advanced Computer Science and Information Systems*, pp. 420–426. <https://doi.org/10.1109/ICACIS.2014.7065883>.
- Ravishankar, H., Prabhu, S.M., Vaidya, V., and Singhal, N. (2016). Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning. In *Proceedings - International Symposium on Biomedical Imaging*, 2016-June, pp. 779–782. <https://doi.org/10.1109/ISBI.2016.7493382>.
- Rawat, V., Jain, A., Shrimali, V., and Rawat, A. (2016). Automatic detection of fetal abnormality using head and abdominal circumference. N.T. Nguyen, Y. Manolopoulos, L. Iliadis, and B. Trawinski, eds. 9876, 525–534. https://doi.org/10.1007/978-3-319-45246-3_50.
- Ryou, H., Yaqub, M., Cavallaro, A., Papageorghiou, A.T., and Alison Noble, J. (2019). Automated 3D ultrasound image analysis for first trimester assessment of fetal health. *Phys. Med. Biol.* 64, 185010. <https://doi.org/10.1088/1361-6560/ab3ad1>.
- Ryou, H., Yaqub, M., Cavallaro, A., Roseman, F., Papageorghiou, A., and Alison Noble, J. (2016). Automated 3D ultrasound biometry planes extraction for first trimester fetal assessment. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10019 LNCS, pp. 196–204. https://doi.org/10.1007/978-3-319-47157-0_24.
- Sahba, F., Tizhoosh, H.R., and Salama, M.M.A. (2006). A reinforcement learning framework for medical image segmentation. In *IEEE International Conference on Neural Networks - Conference Proceedings*, pp. 511–517. <https://doi.org/10.1109/ijcnn.2006.246725>.
- Sahli, H., Sayadi, M., and Rachdi, R. (2020). Intelligent detection of fetal hydrocephalus. *Comput. Methods Biomech. Biomed. Eng.* 8, 641–648. <https://doi.org/10.1080/21681163.2020.1780156>.
- Savadjiev, P., Chong, J., Dohan, A., Vakalopoulou, M., Reinhold, C., Paragios, N., and Gallix, B. (2019). Demystification of AI-driven medical image interpretation: past, present and future. *Eur. Radiol.* 29, 1616–1624. <https://doi.org/10.1007/s00330-018-5674-x>.
- Selvathi, D., and Chandralekha, R. (2021). Fetal biometric based abnormality detection during prenatal development using deep learning techniques. *Multidim. Syst. Sign. Process.* <https://doi.org/10.1007/s11045-021-00765-0>.
- Sen, C. (2017). Preterm labor and preterm birth. *J. Perinat. Med.* 45, 911–913. <https://doi.org/10.1515/jpm-2017-0298>.
- Shahid, N., Rappon, T., and Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: a scoping review. *PLoS One* 14. <https://doi.org/10.1371/journal.pone.0212356>.
- Shiney, O.J., Singh, J.A.P., and Shan, B.P. (2017). A Review on techniques for computer aided diagnosis of soft markers for detection of down syndrome in ultrasound fetal images. *Biomed. Pharmacol. J.* 10, 1559–1568. <https://doi.org/10.13005/bpj/1266>.
- Shuo Wang, B., Ji-Bin, Liu, M., Zhu, M., and John Eisenbrey, P. (2019). Artificial intelligence in ultrasound imaging: current research and applications. *Adv. Ultrasound Diagn. Ther.* 3, 53. <https://doi.org/10.37015/audt.2019.190811>.
- Sinclair, M., Baumgartner, C.F., Matthew, J., Bai, W., Martinez, J.C., Li, Y., Smith, S., Knight, C.L., Kainz, B., Hajnal, J., et al. (2018). Human-level performance on automatic head biometrics in fetal ultrasound using fully convolutional neural networks. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, pp. 714–717. <https://doi.org/10.1109/EMBC.2018.8512278>.
- Singh, T., Kudavelly, S.R., and Suryanarayana, K.V. (2021a). Deep learning based fetal face detection and visualization in prenatal ultrasound. In *Proceedings - International Symposium on Biomedical Imaging*, 2021-April, pp. 1760–1763. <https://doi.org/10.1109/ISBI48211.2021.9433915>.
- Singh, V., Sridar, P., Kim, J., Nanan, R., Poornima, N., Priya, S., Reddy, G.S., Chandrasekaran, S., and Krishnakumar, R. (2021b). Semantic segmentation of cerebellum in 2D fetal ultrasound brain images using convolutional neural networks. *IEEE Access* 9, 85864–85873. <https://doi.org/10.1109/access.2021.3088946>.
- Skeika, E.L., da Luz, M.R., Torres Fernandes, B.J., Siqueira, H.V., and de Andrade, M.L.S.C. (2020). Convolutional neural network to detect and measure fetal skull circumference in ultrasound imaging. *IEEE Access* 8, 191519–191529. <https://doi.org/10.1109/ACCESS.2020.3032376>.
- Sobhaninia, Z., Emami, A., Karimi, N., and Samavi, S. (2020). Localization of fetal head in ultrasound images by multiscale view and deep neural networks. In *25th International Computer Conference (Computer Society of Iran)*, pp. 1–5. <https://doi.org/10.1109/CSICC49403.2020.9050094>.
- Sobhaninia, Z., Rafiei, S., Emami, A., Karimi, N., Najarian, K., Samavi, S., and Reza Soroushmehr, S.M. (2019). Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 6545–6548. <https://doi.org/10.1109/EMBC.2019.8856981>.
- Sridar, P., Kumar, A., Quinton, A., Nanan, R., Kim, J., and Krishnakumar, R. (2019). Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks. *Ultrasound Med. Biol.* 45, 1259–1273. <https://doi.org/10.1016/j.ultrasmedbio.2018.11.016>.
- Sundaresan, V., Bridge, C.P., Ioannou, C., and Noble, J.A. (2017). Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks. In *Proceedings - International Symposium on Biomedical Imaging*, pp. 671–674. <https://doi.org/10.1109/ISBI.2017.7950609>.
- Tan, J., Au, A., Meng, Q., FinesilverSmith, S., Simpson, J., Rueckert, D., Razavi, R., Day, T., Lloyd, D., and Kainz, B. (2020). Automated detection of congenital heart disease in fetal ultrasound screening. *Lect. Notes Comput. Sci.* 12437, 243–252. https://doi.org/10.1007/978-3-030-60334-2_24.
- Tang, X. (2020). The role of artificial intelligence in medical imaging research. *BJR Open* 2, 20190031. <https://doi.org/10.1259/bjro.20190031>.
- Torrents-Barrena, J., Piella, G., Masoller, N., Gratacós, E., Eixarch, E., Ceresa, M., and Ballester, M.Á.G. (2019). Segmentation and classification in MRI and US fetal imaging: recent trends and future prospects. *Med. Image Anal.* 51, 61–88. <https://doi.org/10.1016/j.media.2018.10.003>.
- Toussaint, N., Khanal, B., Sinclair, M., Gomez, A., Skelton, E., Matthew, J., and Schnabel, J.A. (2018). Weakly supervised localisation for fetal ultrasound images. *Lect. Notes Comput. Sci.* 11045, 192–200. https://doi.org/10.1007/978-3-030-00889-5_22.
- van den Heuvel, T.L.A., de Bruijn, D., de Korte, C.L., and van Ginneken, B. (2018). Automated measurement of fetal head circumference using 2D ultrasound images. *PLoS One* 13. <https://doi.org/10.1371/journal.pone.0200412>.
- Wang, S., Hua, Y., Cao, Y., Song, T., Xue, Z., Gong, X., Wang, G., Ma, R., and Guan, H. (2019). Deep learning based fetal middle cerebral Artery segmentation in Large-scale ultrasound images. In *Proceedings - 2018 IEEE International*

- Conference on Bioinformatics and Biomedicine, BIBM2018, pp. 532–539. <https://doi.org/10.1109/BIBM.2018.8621510>.
- Wang, W., Liang, D., Chen, Q., Iwamoto, Y., Han, X.H., Zhang, Q., Hu, H., Lin, L., and Chen, Y.W. (2020). Medical image classification using deep learning. *Intell. Syst. Ref. Libr.* 171, 33–51. https://doi.org/10.1007/978-3-030-32606-7_3.
- Wang, X., Liu, Z., Du, Y., Diao, Y., Liu, P., Lv, G., and Zhang, H. (2021). Recognition of fetal facial ultrasound standard plane based on texture feature fusion. *Comput. Math. Methods Med.* <https://doi.org/10.1155/2021/6656942>.
- Wee, L.K., Min, T.Y., Arooj, A., and Supriyanto, E. (2010). Nuchal translucency marker detection based on artificial neural network and measurement via bidirectional iteration forward propagation. *WSEAS Trans. Inf. Sci. Appl.* 7, 1025–1036.
- Weerasinghe, N.H., Lovell, N.H., Welsh, A.W., and Stevenson, G.N. (2021). Multi-parametric fusion of 3D power Doppler ultrasound for fetal kidney segmentation using fully convolutional neural networks. *IEEE J. Biomed. Health Inform.* 25, 2050–2057. <https://doi.org/10.1109/JBHI.2020.3027318>.
- Whitworth, M., Bricker, L., and Mullan, C. (2015). Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst. Rev.* 2015. <https://doi.org/10.1002/14651858.CD007058.pub3>.
- Wu, L., Cheng, J.Z., Li, S., Lei, B., Wang, T., and Ni, D. (2017). FUIQA: fetal ultrasound image quality assessment with deep convolutional networks. *IEEE Trans. Cybern.* 47, 1336–1349. <https://doi.org/10.1109/TCYB.2017.2671898>.
- Xie, B., Lei, T., Wang, N., Cai, H., Xian, J., He, M., Zhang, L., and Xie, H. (2020a). Computer-aided diagnosis for fetal brain ultrasound images using deep convolutional neural networks. *Int. J. Comput. Assist. Radiol. Surg.* 15, 1303–1312. <https://doi.org/10.1007/s11548-020-02182-3>.
- Xie, H.N., Wang, N., He, M., Zhang, L.H., Cai, H.M., Xian, J.B., Lin, M.F., Zheng, J., and Yang, Y.Z. (2020b). Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. *Ultrasound Obstet. Gynecol.* 56, 579–587. <https://doi.org/10.1002/uog.21967>.
- Xu, L., Gao, S., Shi, L., Wei, B., Liu, X., Zhang, J., and He, Y. (2021). Exploiting vector attention and context prior for ultrasound image segmentation. *Neurocomputing* 454, 461–473. <https://doi.org/10.1016/j.neucom.2021.05.033>.
- Xu, L., Liu, M., Shen, Z., Wang, H., Liu, X., Wang, X., Wang, S., Li, T., Yu, S., Hou, M., et al. (2020a). DW-Net: a cascaded convolutional neural network for apical four-chamber view segmentation in fetal echocardiography. *Comput. Med. Imag. Graph.* 80, 101690. <https://doi.org/10.1016/j.compmedimag.2019.101690>.
- Xu, L., Liu, M., Zhang, J., and He, Y. (2020b). Convolutional-neural-network-based approach for segmentation of apical four-chamber view from fetal echocardiography. *IEEE Access* 8, 80437–80446. <https://doi.org/10.1109/ACCESS.2020.2984630>.
- Yang, T., Han, J., Zhu, H., Li, T., Liu, X., Gu, X., Liu, X., An, S., Zhang, Y., Zhang, Y., and He, Y. (2020a). Segmentation of five components in four chamber view of fetal echocardiography. In *Proceedings - International Symposium on Biomedical Imaging*, 2020-April, pp. 1962–1965. <https://doi.org/10.1109/ISBI45749.2020.9098726>.
- Yang, X., Dou, H., Huang, R., Xue, W., Huang, Y., Qian, J., Zhang, Y., Luo, H., Guo, H., Wang, T., et al. (2021a). Agent with warm start and adaptive dynamic termination for plane localization in 3D ultrasound. *IEEE Trans. Med. Imag.* 11768, 290–298. <https://doi.org/10.1109/TMI.2021.3069663>.
- Yang, X., Huang, Y., Huang, R., Dou, H., Li, R., Qian, J., Huang, X., Shi, W., Chen, C., Zhang, Y., et al. (2021b). Searching collaborative agents for multi-plane localization in 3D ultrasound. *Med. Image Anal.* 72, 102119. <https://doi.org/10.1016/j.media.2021.102119>.
- Yang, X., Ni, D., Qin, J., Li, S., Wang, T., Chen, S., and Heng, P.A. (2014). Standard plane localization in ultrasound by radial component. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 1180–1183. <https://doi.org/10.1109/isbi.2014.6868086>.
- Yang, X., Wang, X., Wang, Y., Dou, H., Li, S., Wen, H., Lin, Y., Heng, P.A., and Ni, D. (2020b). Hybrid attention for automatic segmentation of whole fetal head in prenatal ultrasound volumes. *Comput. Methods Progr. Biomed.* 194, 105519. <https://doi.org/10.1016/j.cmpb.2020.105519>.
- Yang, X., Yu, L., Li, S., Wen, H., Luo, D., Bian, C., Qin, J., Ni, D., and Heng, P.A. (2019). Towards automated semantic segmentation in prenatal volumetric ultrasound. *IEEE Trans. Med. Imag.* 38, 180–193. <https://doi.org/10.1109/TMI.2018.2858779>.
- Yaqub, M., Cuingnet, R., Napolitano, R., Roundhill, D., Papageorghiou, A., Ardon, R., and Noble, J.A. (2013). Volumetric segmentation of key fetal brain structures in 3D ultrasound. *Lect. Notes Comput. Sci.* 8184, 25–32. https://doi.org/10.1007/978-3-319-02267-3_4.
- Yaqub, M., Kelly, B., Papageorghiou, A.T., and Noble, J.A. (2015). Guided Random Forests for identification of key fetal anatomy and image categorization in ultrasound scans. *Lect. Notes Comput. Sci.* 9351, 687–694. https://doi.org/10.1007/978-3-319-24574-4_82.
- Yasaka, K., and Abe, O. (2018). Deep learning and artificial intelligence in radiology: current applications and future directions. *PLoS Med.* 15. <https://doi.org/10.1371/journal.pmed.1002707>.
- Yekdast, R. (2019). An intelligent method for down syndrome detection in fetuses using ultrasound images and deep learning neural networks. *Comput. Res. Prog. Appl. Sci. Eng.* 5, 92–97.
- Yu, Z., Ni, D., Chen, S., Li, S., Wang, T., and Lei, B. (2016). Fetal facial standard plane recognition via very deep convolutional networks. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016-October, pp. 627–630. <https://doi.org/10.1109/EMBC.2016.7590780>.
- Yu, Z., Tan, E.L., Ni, D., Qin, J., Chen, S., Li, S., Lei, B., and Wang, T. (2018). A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition. *IEEE J. Biomed. Health Inform.* 22, 874–885. <https://doi.org/10.1109/JBHI.2017.2705031>.
- Yushkevich, P., Piven, J., Cody, H., and Ho, S. (2006). User-guided level set segmentation of anatomical structures with ITK-SNAP. *Neuroimage* 31, 1116–1128. https://www.academia.edu/download/41218070/User-Guided_Level_Set_Segmentation_of_An20160113-9594-xfzom3.pdf20160115-19908-1771q5v.pdf.
- Zeng, Y., Tsui, P.H., Wu, W., Zhou, Z., and Wu, S. (2021). Fetal ultrasound image segmentation for automatic head circumference biometry using deeply supervised attention-Gated V-Net. *J. Digit. Imag.* 34, 134–148. <https://doi.org/10.1007/s10278-020-00410-5>.
- Zhang, J., Petitjean, C., Lopez, P., and Ainouz, S. (2020a). Direct estimation of fetal head circumference from ultrasound images based on regression CNN. *Proc. Mach. Learn. Res.* 121, 914–922. <http://proceedings.mlr.press/v121/zhang20a.html>.
- Zhang, L., Zhang, J., Li, Z., and Song, Y. (2020b). A multiple-channel and atrous convolution network for ultrasound image segmentation. *Med. Phys.* 47, 6270–6285. <https://doi.org/10.1002/mp.14512>.
- Zhou, S.K., Le, H.N., Luu, K., V Nguyen, H., and Ayache, N. (2021). Deep reinforcement learning in medical imaging: a literature review. *Med. Image Anal.* 73, 102193. <https://doi.org/10.1016/j.media.2021.102193>.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|----------------------------|-------------------------------|---|
| Deposited data | | |
| Studies' methodologies | Contained in the article | N/A |
| Other | | |
| Publicly available dataset | Grand Challenge Fetal dataset | Automated measurement of fetal head circumference using 2D ultrasound images Zenodo |
| | Fetal Planes Dataset | FETAL_PLANES_DB: Common maternal-fetal ultrasound images Zenodo |

RESOURCE AVAILABILITY

Lead contact

Further requests for resources and materials should be directed to and will be fulfilled by the lead contact, Dr. Mowafa Househ (mhouseh@hbku.edu.qa).

Materials availability

This study did not yield new unique reagents.

Data and code availability

- This paper analyzes existing, publicly available data that can be shared by the [lead contact](#) upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

METHODS DETAILS

Eligibility criteria

Studies published in the last eleven years reporting on 'fetus' or 'fetal' were included in this survey. The population of interest was pregnant women in the first, second, and third trimesters with a specific focus on fetal development during these time periods. We included computer vision AI interventions that explore the fetal development process during the prenatal period. Ultrasound image was the only medical image addressed in this survey with different synonyms (e.g., sonography, ultrasonography, echography, and echocardiography).

Search strategy

The bibliographic databases used in this study were PubMed, EMBASE, PsycINFO, IEEE Xplore, ACM Digital Library, ScienceDirect, and Web of Science. Google Scholar was also used as a search engine. The first 200 results were filtered from Google Scholar, which returned a significant number of papers sorted by relevancy to the search subject (20 pages). The search began on June 22, 2021, and ended on June 23, 2021. The original query was broad: we searched terms in all fields in all databases; because of the huge number of irrelevant studies while searching in Web of Science and ScienceDirect, we only searched article title, abstract, and keywords.

A literature retrieval strategy for AI for fetal monitoring

| Databases | Search terms |
|-----------|---|
| PubMed | ("Ultrasound" OR "sonographic" OR "neurosonogram" OR "Sonography" OR "Obstetric" AND (y_10[Filter])) AND ("artificial intelligence" OR "machine learning" OR "neural network" OR "Deep learning" AND (y_10[Filter])) AND ("Fetus" OR "fetal" OR "embryo" OR "baby" AND (y_10[Filter])) Default full text. |

(Continued on next page)

Continued

| Databases | Search terms |
|---------------------|---|
| Embase | ("Ultrasound" OR "sonographic" OR "neurosonogram" OR "Sonography" OR "Obstetric" AND (y_10[Filter])) AND ("artificial intelligence" OR "machine learning" OR "neural network" OR "Deep learning" AND (y_10[Filter])) AND ("Fetus" OR "fetal" OR "embryo" OR "baby" AND (y_10[Filter])) Default full text. |
| PsycINFO | ("Ultrasound" OR "sonographic" OR "neurosonogram" OR "Sonography" OR "Obstetric" AND (y_10[Filter])) AND ("artificial intelligence" OR "machine learning" OR "neural network" OR "Deep learning" AND (y_10[Filter])) AND ("Fetus" OR "fetal" OR "embryo" OR "baby" AND (y_10[Filter])) full text. |
| ScienceDirect | ("Fetus" OR "fetal") AND ("artificial intelligence" OR "neural network") AND ("Ultrasound" OR "sonography") |
| IEEE | "Fetus" OR "All Metadata": "fetal" OR "All Metadata": "embryo" OR "All Metadata": "baby" AND ("All Metadata": "artificial intelligence" OR "All Metadata": "machine learning" OR "All Metadata": "neural network" OR "All Metadata": "Deep learning") AND ("All Metadata": "Ultrasound" OR "All Metadata": "sonographic" OR "All Metadata": "neurosonogram" OR "All Metadata": "Sonography" OR "All Metadata": "Obstetric") Filters Applied: 2010 - 2021. |
| ACM Digital library | [[All: "fetus" OR All: "fetal" OR All: "embryo" OR All: "baby"]] AND [[All: "artificial intelligence" OR All: "machine learning" OR All: "neural network" OR All: "deep learning"]] AND [[All: "ultrasound" OR All: "sonographic" OR All: "neurosonogram" OR All: "sonography" OR All: "obstetric"]] AND [Publication Date: (01/01/2010 TO 06/30/2021)] |
| Google Scholar | ("Fetus" OR "fetal" OR "embryo" OR "baby") AND ("artificial intelligence" OR "machine learning" OR "neural network" OR "Deep learning") AND ("Ultrasound" OR "sonographic" OR "neurosonogram" OR "Sonography" OR "Obstetric") |
| Web of science | ((ALL=("Fetus" OR "fetal" OR "embryo" OR "baby")) AND ALL=("artificial intelligence" OR "machine learning" OR "neural network" OR "Deep learning")) AND ALL=("Ultrasound" OR "sonographic" OR "neurosonograms" OR "Sonography" OR "Obstetric"). |

Evaluating tool and reporting standard

This study used a few tools for quality evaluation and reporting standards. First, we used Raya web-based software for study selection, including duplicate remover, title, and abstract screening. Also, the Rayyan software helped us label each study for fetal and AI task categorization in further steps. Second, we used an excel sheet for data extraction for each article. The data extraction form is seen in [Table S1](#). Third, after we finalized the total number of selected studies, we used bibliometrics analysis software to screen them and ensure that we did not miss any relevant studies. Finally, we evaluated each study independently, including the study method, dataset, novelty, and result.

QUANTIFICATION AND STATISTICAL ANALYSIS

This paper evaluates the statistical and quantitative analytic methods used in published studies. The authors of this paper did not conduct additional quantified analysis, such as meta-analysis.