Medical Image Analysis 51 (2019) 61–88

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Segmentation and classification in MRI and US fetal imaging: Recent trends and future prospects☆

Jordina Torrents-Barrena[a,*], Gemma Piella[a], Narcís Masoller[b,c], Eduard Gratacós[b,c], Elisenda Eixarch[b,c], Mario Ceresa[a], Miguel Ángel González Ballester[a,d]

[a] BCN MedTech, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
[b] BCNatal - Barcelona Center for Maternal-Fetal and Neonatal Medicine (Hospital Clínic and Hospital Sant Joan de Déu), IDIBAPS, University of Barcelona, Barcelona, Spain
[c] Center for Biomedical Research on Rare Diseases (CIBER-ER), Barcelona, Spain
[d] ICREA, Barcelona, Spain

ARTICLE INFO

Article history:
Received 9 December 2017
Revised 9 October 2018
Accepted 18 October 2018
Available online 19 October 2018

Keywords:
Fetal magnetic resonance imaging
Fetal (3D) ultrasound
Fetal structural abnormalities
Placenta
Heart
Lungs
Liver
Segmentation
Classification
In utero image analysis

ABSTRACT

Fetal imaging is a burgeoning topic. New advancements in both magnetic resonance imaging and (3D) ultrasound currently allow doctors to diagnose fetal structural abnormalities such as those involved in twin-to-twin transfusion syndrome, gestational diabetes mellitus, pulmonary sequestration and hypoplasia, congenital heart disease, diaphragmatic hernia, ventriculomegaly, etc. Considering the continued breakthroughs *in utero* image analysis and (3D) reconstruction models, it is now possible to gain more insight into the ongoing development of the fetus. Best prenatal diagnosis performances rely on the conscious preparation of the clinicians in terms of fetal anatomy knowledge. Therefore, fetal imaging will likely span and increase its prevalence in the forthcoming years. This review covers state-of-the-art segmentation and classification methodologies for the whole fetus and, more specifically, the fetal brain, lungs, liver, heart and placenta in magnetic resonance imaging and (3D) ultrasound for the first time. Potential applications of the aforementioned methods into clinical settings are also inspected. Finally, improvements in existing approaches as well as most promising avenues to new areas of research are briefly outlined.

1. Introduction

Prenatal imaging technology for fetal diagnosis has rapidly evolved (Yan (2017)). Two-dimensional ultrasound (2D US) is the primary screening modality for pregnancy evaluation because of its relative low cost, real-time imaging, lack of harmful effects to both fetus and mother, and high resolution. Three- and four -dimensional ultrasound (3D / 4D US) with additional sonography modalities such as color / power Doppler are increasingly available and have been successfully employed to detect fetal structural abnormalities (Roy-Lacroix et al. (2017a)).

Although US is still the standard tool for fetal imaging, it provides a small field-of-view, limited soft-tissue acoustic contrast, beam attenuation by adipose tissue, and poor image quality. Furthermore, important limitations arise in the context of an intrauterine environment, such as operator variability, fetal position, effects due to gestational age (poor visualization, skull ossification), and fetal tissue definition (Roy-Lacroix et al. (2017a); Reddy et al. (2008)). In cases where abnormalities cannot be adequately assessed with US, fetal magnetic resonance imaging (MRI) can be a good alternative. Early studies using MRI in the evaluation of fetal morphology were hindered due to image degradation by fetal motion on MRI standard sequences and the relatively long acquisition time. Nevertheless, with ultrafast MRI techniques such as single-shot fast-spin echo and half-Fourier acquisition turbo spin echo (HASTE), fetal motion artifacts are minimized. Dramatic reductions in image acquisition time and improved acquisition protocols for multi-planar imaging have enabled MRI to be used in fetal evaluation (Story and Rutherford (2015)). High-quality fetal images are now routinely achieved without maternal or fetal sedation. However, limitations of fetal MRI include the need of equipment availability and radiology expertise, higher cost, and longer time to perform a complete examination.

The increasing advancement and use of fetal imaging for computer-assisted prenatal organ assessment provides essential

☆ *Conflict of interest:* There are no conflicts of interest.

* Corresponding author.
E-mail address: jordina.torrents@upf.edu (J. Torrents-Barrena).

https://doi.org/10.1016/j.media.2018.10.003
1361-8415/

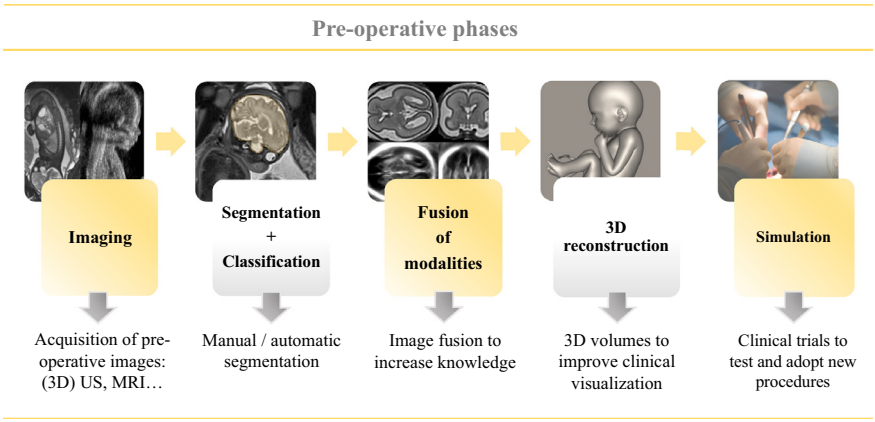

Fig. 1. Steps involved in the pre-operative surgical planning.

clinical feedback to design improved ad hoc surgical frameworks. A better understanding of fetal pathophysiology and early diagnosis has paved the way for the development of *in utero* therapies rather than postnatally (Maselli and Badillo (2016)). In this way, (3D) US and MRI provide crucial anatomic information that can be helpful in planning antenatal care and surgical procedures, as well as predicting fetal development and outcome. Fetal surgery has emerged as a multidisciplinary field capable of improving fetus outcomes for a wide range of interventions, such as laser treatment for twin-to-twin transfusion syndrome (TTTS) and open fetal surgery for spina bifida (Pratt et al. (2015); Slaghekke et al. (2014); Adzick et al. (2011)). Fetal interventions are expected to increase in breadth and prevalence as novel surgical or computational procedures demonstrate to be clinically valid (Pratt et al. (2015); Mathis et al. (2015)). As in any surgery, best results are achieved when doctors are prepared pre-operatively with a detailed understanding of the baby's anatomy thanks to fetus-specific 3D visualizations obtained from medical images.

The pre-operative phase in fetal surgery involves the following steps (see Fig. 1) (Pratt et al. (2015)). First, images of the target structures are taken. Segmentation and classification methods are then applied to distinguish between the regions of interest and the background organs and tissues that can be discarded. The fusion of several image modalities allows completing the target structures from the different partial image views. Next, volume rendering algorithms are applied in the reconstruction stage to yield a 3D model in which the structural relationships are disclosed. In the simulation phase, surgeons benefit from this model to visualize the anatomical structures from different perspectives, and to understand their intricate relationships.

Moreover, segmentation and classification techniques can also provide imaging biomarkers for the prediction of fetal development and outcome. The shape, volume, morphometry and texture of fetal organs can be characterized during pregnancy to assess fetal health and predict complicated pregnancies (Dahdouh et al. (2018)). Early biomarkers of fetal organs disease that may impair fetal growth and well-being open up brand-new opportunities to intervene and protect vulnerable fetuses.

This review attempts to fill existing gaps in the literature. We sought to uncover both the literature in fetal imaging and to illustrate the usefulness of novel approaches to provide additional and improved diagnostic information. To the best of our knowledge, the state-of-the-art of segmentation and classification methods for fetal organs such as placenta, lungs, liver and heart, or even the whole fetus, in (3D) US and MRI have not been reviewed so far. Considering the relevance and the steady progress made in the field, we deemed it necessary to survey them in detail. For completeness, fetal brain is also covered. While several reviews of fetal brain MRI exist (see Section 3.2.1), considerably less literature can be found on fetal brain US. To our knowledge, this is the first review that addresses it.

The contents are structured as follows. Section 2 describes the search criteria for this work. Section 3 reviews the main works on fetal MRI and 2D/3D US imaging, categorized by fetal organs: the whole fetus, brain, placenta, lungs, liver, thorax, and heart. Section 4 provides examples of fetal surgery applications for which the methods reviewed in Section 3 can be employed at the pre-operative phase. Section 5 briefly overviews the validation performed on the state-of-the-art reviewed. Lastly, Section 6 discusses the most useful contributions to the field by highlighting some cutting-edge ideas.

2. Method

A systematic review on the use of segmentation and classification methods in fetal imaging has been performed. To scrutinize the literature, medical subject headings (*MeSH*)[1] and keywords related to the scope of this review (segmentation, classification, computer vision, machine learning, fetal MRI and (3D) US imaging) were combined with medical terms associated to prenatal diagnosis. *PUBMED*[2], *EMBASE*[3] and *ISI Web of Knowledge*[4] platforms were searched electronically on September 2017 and updated on July 2018. Reference lists of relevant articles and reviews were hand searched for additional literature, and the papers that cited them reviewed. Studies published at any time that discussed the use of MRI and (3D) US segmentation and classification methods specifically for fetal organs such as the whole fetus, brain, placenta, lungs, liver and heart were included. A total of 321 studies were identified, of which 189 were excluded after review of the title and abstract. The remaining 132 were read in full, and 9 excluded to be out of the scope. The works included in this review describe cases in which scientific algorithms could be used to identify fetal structural abnormalities such as those involved in gestational diabetes mellitus, twin-to-twin transfusion syndrome, pulmonary sequestration and hypoplasia, congenital diaphragmatic hernia, and congenital heart disease, etc. Notice that the here-reviewed studies are of purely scientific type, since none of the technologies have been marketed.

[1] MeSH: https://meshb.nlm.nih.gov/search.
[2] PUBMED: www.ncbi.nlm.nih.gov/pubmed/.
[3] EMBASE: www.elsevier.com/solutions/embase-biomedical-research.
[4] ISI Web of Knowledge: apps.webofknowledge.com/.

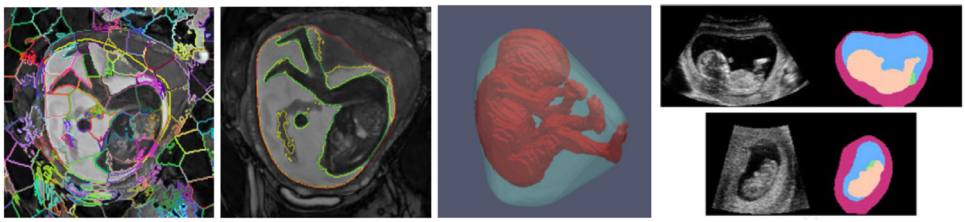

Fig. 2. Figure adapted from *Zhang et al. (2016)* and *Anquez et al. (2013)*: Whole fetus segmentation in MRI (*left*) and US (*right*).

3. Categorization by fetal organ

The main challenge to automatically localize and classify the fetal organs such as brain, placenta, heart, lungs and liver, or even the whole fetus is the unpredictable fetus position and arbitrary orientation. This section gathers 123 articles that show quantitative / qualitative and robust validation of their methods, these have been assessed clinically by doctors and have undergone a peer-review process.

3.1. Whole fetus

Fetal body segmentation and volumetry are studied in antenatal research and clinical assessments for a wide range of applications. Typically, these measurements still rely on manual and subjective clinical tracing. Effective and automatic monitoring of fetal growth, body volume, weight and early diagnosis of premature anomalies requires fetal imaging and analysis across gestation (see Fig. 2). Compared with obstetrical US, which routinely covers pregnancy stage determination, placenta positioning, fetal growth, quantitative measurement of biometric markers, or characterization of potential pathologies, among others, MRI presents an extended field-of-view and superior soft-tissue contrast. MRI has thus become popular for prenatal examination due to the limited field-of-view of fetal US after the first trimester.

3.1.1. Whole fetus segmentation on fetal MRI

A semi-automatic method was proposed by Anquez et al. (2010) to segment the fetal envelope on ante-natal MRI acquired using the *Steady State Free Precession* sequence. A simplified skeleton was firstly extracted, identifying a group of 25 landmark points. An articulated model of a generic fetus was then registered to the aforementioned skeleton. This registration was employed to define a narrow-band containing the fetal envelope, which was segmented via graph-cuts. Although authors obtained a mean distance of 1.4 mm and a Cohen's kappa of 0.89, prior knowledge such as contrast, shape modeling and biometric information could be exploited to guide, and consequently, improve the segmentation. The same authors (Bibin et al. (2009)) also implemented a computational framework to generate synthetic pregnant women through embryos and women segmented from 3D US (from 8 to 13 gestational weeks (GW)) and MRI (from 30 to 34 GW), respectively. The method included image segmentation, contour regularization, mesh-based surface reconstruction, and model integration. All models were validated by obstetricians and pediatric radiographers.

Zhang et al. (2016) created an automatic graph-based approach that over-segmented 10 uterine images to classify the amniotic fluid, fetal body and background. Once the amniotic fluid was detected, a specific ROI was defined where a set of edge nodes were randomly selected as possible targets. After ranking all graph nodes, the fetal body was delineated iteratively achieving a Dice coefficient (DC) of 0.69 ± 0.08.

A distinct approach was realized by Zito et al. (2013) in 20 singleton pregnancies. The fetal body was determined on each MRI slice through free-form ROIs. The corresponding fetal weight was also calculated using Baker's equation and compared to actual birth weight. The relative error 1.96% obtained was mainly associated with patient's body mass index and not with the type of MRI sequence used. The use of thicker slices (< 6 or 4 mm) significantly reduced the time for fetal body segmentation.

3.1.2. Whole fetus segmentation on fetal (3D) US

One of the first works was proposed by Gupta et al. (2011). They used a fuzzy C-means (FCM) clustering with wavelet-based texture features and Support Vector Machines (SVM) for initial label prediction. Afterwards, a conditional random field (CRF) was employed to exploit context information and further refine the fetal body segmentation. The proposed methodology was tested qualitatively on 2 images. Krishnan et al. (2014) used anisotropic diffusion to enhance the fetus edges. Both isointensity and edge-focusing contour approaches were employed to determine the fetus shape as 3D points. After fitting the head and body, the volume was calculated with an error of 4.8%.

The approach implemented by Anquez et al. (2013) used a Bayesian formulation to partition the US into amniotic fluid and fetal tissues. An energy functional was minimized through a level set to identify the optimal partition. They combined Rayleigh, normal, exponential, and gamma distribution models to compute the region homogeneity constraints. The method, validated on 19 antenatal images, was shown to be robust (DC of 0.89) to different initialization types. A variational formulation of the level set algorithm was applied by Choong et al. (2012). An internal energy penalized the deviation of the level set from a signed distance function, while an external energy enforced the motion of the zero level set towards the desired image features. This variational level set presented three advantages: 1) a larger time step was utilized to speed up the curve evolution, 2) the re-initialization was eliminated, and 3) the level set curve was implemented for different object shapes. The total difference percentage obtained for fetal body extraction was around 2.38%. Another variational shape-guided method for extracting the fetus envelope was presented by Dahdouh et al. (2013, 2015). To compensate the lack of contrast and explicit boundaries, authors' method combined the pixel intensity distribution, with a fetal envelope shape prior and a back model depending on fetus age. The intensity distributions and the shape prior were encoded using translation-invariant *Legendre* moments and added as energy terms in the functional to be optimized. The back model was only employed during the post-processing phase. The method was trained and tested using 5 and 11 US volumes between 8–13 GW. Mean values of 0.72, 0.89 and 0.97 were obtained for similarity, sensitivity (SEN) and specificity (SPE), respectively. However, the fetus was under-segmented and some false positives still remained due to maternal tissues, and the back model did not perfectly match the fetus as it was generic.

Li et al. (2017) applied a deep learning architecture to segment the amniotic fluid and fetal tissues. The model encoded and decoded the input US into down-scaled and up-scaled feature maps using several convolution and pooling structures. Additional 1×1 convolution layers were adopted to discriminate and enhance

the feature representations. Cross-domain transfer learning was employed to avoid the need of training with a large dataset. As a result, they achieved an accuracy (ACC) of 93%, 67% and 78% for background, fetal body and amniotic fluid, respectively. Yang et al. (2017b) simultaneously segmented the fetus, gestational sac and placenta. Based on their customized 3D Fully Convolutional Network (CNN), a Recurrent Neural Network (RNN) was applied to explore 3D semantic knowledge from a sequential perspective, and refine the local segmentation corrupted by US ubiquitous boundary uncertainties. Finally, a hierarchical deep supervision mechanism was also introduced to boost the information flow within the RNN. DC of 0.88, 0.89 and 0.64 were obtained for fetus, gestational sac and placenta, respectively.

3.1.3. Whole fetus standard US plane classification

Precise acquisition of standard US planes of the whole fetus (*e.g.,* fetal abdominal, face axial and four-chamber view) is a critical step in obstetric diagnosis. The main goal is to confirm fetal viability, establish gestational age, and discover prenatal malformations. Clinically, US standard plane is manually acquired by finding the view with concurrent presence of fetal key anatomical structures. This procedure heavily relies on doctors' expertise and is very challenging due to the high intra-class variations of US standard planes resulting from acoustic shadows, deformations of soft tissues and various transducer orientations. Therefore, automatic identification approaches are required to improve the examination efficiency.

One of the first studies on this domain was presented by Carneiro et al. (2008a), which introduced a probabilistic model to combine discriminative and generative classifiers with contextual information and sequential sampling. The designed system required user queries in the form of semantic keywords to represent structures of interest. Afterwards, the system automatically displayed standardized planes and produced biometric measurements of the fetal anatomies. The framework was capable to solve position, orientation and size of three different fetal anatomies in less than 10 *s*. The validation was performed on 40 US volumes. Later, Rahmatullah et al. (2011) basically used the AdaBoost algorithm to categorize the most probable standard plane slices using the classifier scores. Experimental results on 30 US volumes showed a recall and precision of 91.29% and 76.29%, respectively. Two cascade AdaBoost classifiers were also employed by Zhang et al. (2012) to locate the gestational sac standardized plane by exploiting a coarse-to-fine detection scheme. A database-guided multi-scale normalized cuts algorithm generated the initial contour of the sac. The system was validated on 31 US videos, and obtained a system-radiologist performance difference of 7.5% $\pm$ 5.0%, 5.5% $\pm$ 5.2%, and 6.5% $\pm$ 4.6% for early gestational sac selection, length and depth (diameter) measurements, respectively. Regression random forests (RFs) were used in Yaqub et al. (2014). Constraints such as 1) the use of voxels' information and strength to weight the training objective function, and 2) the introduction of a regularization through a geometrical feature, were implemented to improve the standard plane detection. Results on 87 3D fetal brain US between 23–27 GW indicated that this method was more reproducible than manual plane finding. One year later, Yaqub et al. (2015) guided several RF to categorize the following anatomical US views: head in the trans-ventricular and trans-cerebellar planes, 4-chamber view of the heart, face, abdominal circumference, femur, spine and others such as limbs, kidneys, cord insertion, etc. They extracted translation and orientation invariant features at multiple spatial resolutions from ROIs that contained meaningful fetal structures. Evaluated on approximately 30K US, the guided RF achieved an accuracy of 75%, while traditional RF was 65%. Similarly, both Ni et al. (2014) and Yang et al. (2014) created a radial component-based model to describe the geometric constraints of the fetal stomach bubble, umbilical vein and spine through the fetal abdominal standard plane. A selective search method produced the location for the spine and umbilical vein using the vessel probability map algorithm (Penney et al. (2004)). This map was obtained by calculating the image intensity for each pixel, which quantifies the likeness to the typical vessel appearance of a dark center with a relatively bright rim. Next, a RF classifier was used to find the key anatomical structures at their probable locations within the regions constrained by the radial component-based model. A second-level classifier combined the results from the component detection to differentiate the fetal abdominal standard plane. Experimental results on 223 US videos showed an ACC of 85.6%. The work presented by Ni et al. (2013) employed a similar methodology on 100 fetal abdomen videos, but they used an AdaBoost classifier instead of a RF. Other advanced classifiers such as SVM were applied by Lei et al. (2015a) to recognize fetal facial standard planes. Densely root scale invariant and spatial transform features were extracted and encoded using Fisher vector. Experimental results demonstrated an ACC of 93.27% and a mean average precision of 99.19% on 187, 192 and 203 axial, coronal and sagittal plane images, and 1153 images without any identified standard plane.

Very differently, a two-phase algorithm was presented by Liu et al. (2012) to recognize fetal head scan planes. Cranium detection was followed by template matching to capture the internal head structure used in the clinical recognition process of the planes. Then, an active appearance model was trained on 51 correct standard US planes to know the presence or absence of the inclusions within the head. The final error rate was around 28%.

Predictably, researchers also extended deep neural networks to standard plane classification. Chen et al. (2015b, 2014) firstly trained a CNN using the well-known ImageNet database. The domain transferred CNN was then constructed for fetal abdominal standard plane localization in US videos. Initially, the convolutional layers of the base CNN were implanted into the transferred CNN. Authors performed a fine-tuning of the parameters encoded in the convolutional layers to subsequently train the fully connected layers of the transferred CNN with specific US samples. An ACC and F1-measure of 89.6% and 0.71 were achieved, respectively. Yu et al. (2018) also designed a network with 16 convolutional and 3 fully-connected layers. A global average pooling was adopted to reduce network parameters as well as transfer learning and data augmentation, which alleviated the overfitting and improved the performance. Authors' dataset included 375, 257 and 405 axial, coronal and sagittal plane images, and 3812 images without any identified standard plane. A precision of 96.98% was finally achieved. The study made by Baumgartner et al. (2016) automatically detected 12 standard planes at more than 100 US fps. Their fully CNN was extended to obtain saliency maps highlighting the ROI with the highest prediction outcome in an unsupervised manner. A large dataset of 1003 scans was used to evaluate the method, obtaining a precision and recall of 69% and 80%, respectively. Furthermore, the network was capable to retrospectively retrieve correct scan planes with an ACC of 71% for cardiac views and 81% for non-cardiac views. The work in Chen et al. (2015a) explored spatio-temporal regions through a CNN and a long-short term memory (LSTM) with knowledge transfer across multi-tasks. Firstly, a ROI classifier was trained across multi-tasks to locate the most discriminative US regions for standard plane detection. Secondly, the temporal information of the previous extracted features was consecutively explored using the LSTM. Thirdly, the score of each frame was obtained by averaging all LSTM predictions. Authors collected 631 US videos between 18–40 GW. Standard plane detection of the fetal abdominal, face and four-chamber view of the heart obtained a mean ACC of 88% and a F1-measure of 0.68. Recently, Baumgartner et al. (2017) proposed a two-fold CNN named *SonoNet* that was able to detect 13 fetal standard views and provide a localization of several fetal structures via

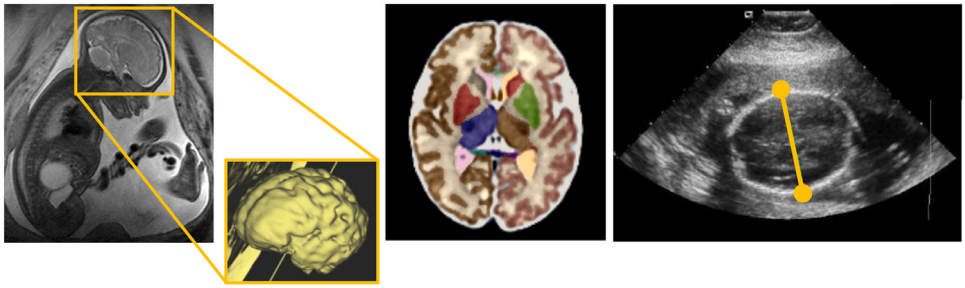

Fig. 3. Detection of the fetal brain: (*left*) 3D model of the brain and categorization of its internal structures (MRI), and (*right*) measurement of the fetal skull (3D US).

bounding boxes. *SonoNet* was inspired by the VGG-16 model (13 convolutional and 3 fully-connected layers). The network performed a real-time isolation of the target anatomy using weak supervision based on image-level labels only. Authors achieved an average F1-score of 0.798 in standard plane detection, an ACC of 90.09% for retrospective frame retrieval, and an ACC of 77.8% on the localization task. Schlemper et al. (2018) introduced self-gated soft-attention mechanisms into the *SonoNet* architecture (Baumgartner et al. (2017)). The attention module provided a fast (weakly-supervised) soft object localization during forward pass, demonstrating that the attended features indeed correlate to the anatomy of interest. In particular, authors implemented a normalization strategy for the attention map, some gridded attention mechanisms, and aggregation strategies. The proposed architecture achieved an ACC and F1 score of 98% and 0.93, respectively. Li et al. (2018a,b) presented an iterative transformation network that learned the relationship between a 2D plane of the brain and the transformation parameters required to move that plane towards the corresponding standard plane. They also introduced additional classification probability outputs to act as confidence measures for the regressed transformation parameters. The effect of two training loss functions to exploit the geometry and the data of the extracted plane was explored. When evaluated on 72 subjects, this approach obtained an error of 3.83 mm and 3.80 mm for the transventricular and transcerebellar planes, respectively.

3.2. Brain

Quantitative methods for fetal brain analysis are essential in clinical decision-making and neuroscience research. Although US imaging is mainly employed for fetal brain examination, some constraints such as reduced amniotic fluid, maternal obesity, improper fetal head position, twin pregnancies, and bony reverberation artifacts from the skull, hamper an accurate assessment (see Fig. 3). Over the past few years, MRI scans have gained a certain momentum to be used for those studies.

3.2.1. Brain segmentation on fetal MRI

Despite advances in the image acquisition, the perinatal brain MRI further exhibit domain-specific challenges such as 1) the increased occurrence of unpredictable motion artifacts and the need of faster and multiple acquisitions to correct them, 2) the lower contrast-to-noise ratio because of the small size of the fetal brain and shorter scanning periods, 3) the exhibition of a reversed gray and white matter contrast compared to the adult MRI data, 4) the large variation in shape and appearance of perinatal brain structures due to the rapid brain development (*i.e.,* the cortical ribbon), and 5) the lack of manually-labelled atlases across several scan ages.

Although fetal brain segmentation in MRI has been reported by Studholme (2011), Studholme and Rousseau (2014), Benkarim et al. (2017) and Makropoulos et al. (2018), we provide a summarized overview. A parametric technique based on a three-stage process was adopted by Anquez et al. (2009). Authors initially identified the fetus eyes via template matching and contrast, morphological operations and biometrical information in the mid-sagittal plane. A shape was then registered to the entire stack to extract a volume of the fetal brain on 24 MRI between 30–35 GW using traditional graph-cuts. A Cohen's kappa measure of 0.98 was finally obtained. Moreover, Ison et al. (2012), Keraudren et al. (2014) and Kainz et al. (2014a) utilized classical computer vision and machine learning techniques. In Ison et al. (2012) analyzed 43 MRI cases aged between 18–30 GW. A RF classifier was used with 3D Haar descriptors to suppress the influence of maternal tissues and detect possible fetal brain locations. This organ was then segmented through a Markov random field (MRF). The results showed ROIs enclosing an average of 88% of the brain with center prediction ACC of 81% in the coronal view, and 78% with center prediction ACC of 60% in the axial and sagittal views. Keraudren et al. (2014) extracted SIFT features from maximally stable extremal regions that were then clustered using K-means. The classification into brain and non-brain was performed through a SVM classifier. Afterwards, this initial segmentation was refined with a RF classifier applied on 2D patches and CRFs. Their approach was assessed on 66 healthy fetuses ranged from 22 to 39 GW, which achieved a DC of 0.93. Kainz et al. (2014a) followed a similar methodology but trained a RF classifier on 3D Gabor descriptors. The segmentation was also refined using a 2D level-set. A final DC of 0.90 was produced.

Atlas-based fusion algorithms have been widely employed for fetal brain localization. Indeed, Taimouri et al. (2015) developed a template-to-slice block matching that found spatiotemporal (4D) atlas correlations of the fetal brain to the corresponding MRI slice. The method was evaluated on 366 MRI scans of 30 women. Final results demonstrated a 94% of success rate. In Wright et al. (2014), spatio-temporal atlas were used as tissue priors in an expectation-maximization approach with second order MRF regularization. This resulted in an improvement of the cortical boundary estimate and an average surface discrepancy of 1 mm. Tourbier et al. (2015) evaluated the performance of multiple atlas fusion strategies with global weighting based on the normalized correlation coefficient. The proposed methodology outperformed other approaches with a DC above 0.90.

Deep learning has been also exploited to overcome the aforementioned challenges. Khalili et al. (2017) employed a 2D multi-scale CNN to learn spatial and detailed local information from larger context. A total of 30 fetal T2-weighted MRI scans were used to train the network, which achieved a DC of 0.91, 0.95 and 0.92 for the axial, coronal and sagittal images, respectively. Similarly, Salehi et al. (2018) developed a fully CNN based on 2D U-Net with auto-context to independently extract 2D sections of the fetal brain in real-time. They compared the presented architecture to two alternative fast methods: a voxel-wise fully CNN, and a method that utilized SIFT features, RF and CRF. The network was trained on 250 MRI stacks, and tested on 17 normal fetuses

and 18 extremely challenging cases. Mean DC of 0.96 and 0.79 for normal and difficult sets were calculated, respectively. Notice that the method proposed by Rajchl et al. (2017) (see Section 3.5.1) also segmented the fetal brain.

3.2.2. Brain segmentation on fetal (3D) US

The fetal head circumference is usually measured on either the trans-thalamic or trans-ventricular US planes and subsequently utilized in dating pregnancies beyond 14 GW. Finding such brain anatomical structures is not straightforward and highly subjective due to the affected visibility caused by many conditions such as the US machine, maternal tissue, fetus orientation, sonographer experience, among others.

In the literature, there are only a handful of published papers based on the detection of the fetal brain and its internal structures in (3D) US. A spherical harmonics model was applied by Velásquez-Rodríguez et al. (2015a,b) to segment the cerebellum. The corresponding triangular mesh was generated for each shape using matching cubes and modeled with the aforementioned spherical harmonic functions. The optimization was based on a gray-level energy function and voxel profiles. A mean DC of 0.75 was achieved on 10 US. Becker et al. (2010) bet on a 3D point distribution model to automatically segment the cerebellum. The model was adjusted to 5 fetal brain 3D US through a genetic algorithm in order to optimize a fitting function. A 3D statistical shape model was also presented by Gutiérrez-Becker et al. (2013) jointly with an ad-hoc objective function optimized via the *Nelder-Mead* simplex algorithm. A total of 20 pregnant women between 18 - 24 GW were evaluated with an intra-class correlation coefficient of 0.85 and a mean DC of 0.8. Moreover, traditional frameworks such as VOCAL were also used by Benavides-Serralde et al. (2009) to detect the intracranial, frontal, thalamic and cerebellar brain structures. Volume measurements were done by an inter-/intra-observer in 78 fetuses between 28–34 GW. The following intra-class correlation coefficients were obtained: intracranial = 0.97, cerebellar = 0.69, frontal = 0.66 and thalamic = 0.54.

Yaqub et al. (2013) built a decision RF framework to segment 5 fetal brain structures. Authors put the volumes into a common coordinate space and employed template deformation to extract the skull. The mid-sagittal plane was approximated using a Hough transform on the responses of a plate detector at different scales. The neck orientation was then utilized to find the fetal head orientation and center through the mid-sagittal plane and eye positions. Some features based on the skull, head center, and left and right eyes centers were integrated within the RF framework to efficiently guide the five-fold segmentation. The following ACCs were achieved in 51 fetuses between 18–26 GW: 79 $\pm$ 9% for the choroid plexus, 82 $\pm$ 10% for the lateral posterior ventricle cavity, 74 $\pm$ 11% for the cavum septum pellucidi, and 63 $\pm$ 15% for the cerebellum. Another RF model was previously implemented by Yaqub et al. (2012) and Napolitano et al. (2012), which implicitly integrated Haar, cuboid, binary and unary features to be classified into the aforementioned classes. In this case, 40 fetal brain US between 19–24 GW were used. The best ACCs were reported in Yaqub et al. (2012) for choroid plexus (92.9%), posterior ventricle cavity (91.1%), cavum septum pellucidum (91.9%), and cerebellum (91.9%). Recently, Cerrolaza et al. (2017) designed another RF-based segmentation method to efficiently integrate semantic and structural information via generalized geodesic distance transform and contextual trees. A structured label space was finally employed to capture the morphological variability of the fetal skull. An averaged DC of 0.80 $\pm$ 0.03 was obtained after 3 *h* of training. Carneiro et al. (2008b) trained a sequence of discriminative AdaBoost classifiers. The system required the image and the measurement of the fetal body and heat to finally segment the rectangular ROI containing such fetal anatomies (it took less than 0.5*s*). A total of 1426 and 325 training samples for the head and fetal body were collected, respectively.

Differently, the approach designed in Cuingnet et al. (2013) aided the navigation through the brain by aligning the fetal head in less than 1 *s*. Such alignment was computed using a reference frame based on the skull, the mid-sagittal plane and the eyes thanks to their stability across different acquisitions. An ellipsoidal shell model was previously needed to segment the fetal head. The skull segmentation was performed using this model as a deformable template. Maximal alignment errors ranged from 5.1 to 5.8 mm on a dataset of 78 3D US between 19–24 GW.

3.3. Placenta

The placenta is an organ that maintains the fetus and the uterine wall connected to allow nutrient uptake, thermo-regulation, respiratory exchange via the mother's blood, and transfer of waste products. It also prevents against infections and segregates hormones to support pregnancy. Pathologic conditions of the placenta include circulation abnormalities as in TTTS, which produce an unequal blood distribution and a fatal outcome to one or both twins. Also, an abnormally adherent placenta (accreta) invading the myometrium can lead to an increment of heavy bleeding during delivery. Placental insufficiency is known as the inability to provide nutrients and remove harmful waste. If it happens for a long period of time, it may lead to intrauterine growth restriction (IUGR). The placenta is the only biological structure that can cause the death or injury of more than one person at the same time. Therefore, this organ plays a critical role in fetus health during gestation as its examination may yield information about maternal or fetal disorders. In this context, different clinical assessments have been performed to quantify the impact of placenta insufficiency on fetus healthcare and outcome (Baschat (2004); Luo et al. (2017); Lean et al. (2017); Phillips et al. (2017)).

Quantitative assessment of the human placenta in MRI and US requires a precise segmentation and classification, which is challenging due to the high variability of its position, orientation, thickness, shape and appearance. 3D data acquisition of the entire uterus is also corrupted by motion artifacts between individual slices because of both fetal movements and maternal respiratory motion. The overall anatomy is therefore displaced. Only few works are reported in the literature, thus no established procedure for assessing the placenta condition is currently available.

3.3.1. Placenta segmentation on fetal MRI

Placenta MRI acquisition provides multi-planar imaging as well as a large field of view. Indeed, placenta segmentation is improved in terms of accuracy and robustness by acquiring several MRI stacks oriented differently (axial, sagittal and coronal). Such stacks (see Fig. 4) can be utilized simultaneously with their complementary resolution in several directions to support placenta co-segmentation. Moreover, the large amount of amniotic fluid leads to a strong field attenuation due to the conductivity artifact and the low image quality. The pixel intensity change between the amniotic fluid and both the mother and fetal subcutaneous fat and tissues also facilitates the placenta extraction.

The first semi-automated approach for placenta segmentation in fetal MRI, termed Slic-Seg, was reported by Wang et al. (2015a) and Wang et al. (2015b). An interactive machine learning based method which minimizes user interaction was implemented. The user needs only to select a starting slice, and to draw a few scribbles to indicate foreground and background. For each pixel, the mean and standard deviation, as well as texture features obtained by gray level co-occurrence matrix (GLCM) and Haar wavelet, were extracted. From these input features, a RF learned to predict the probability of an unlabeled

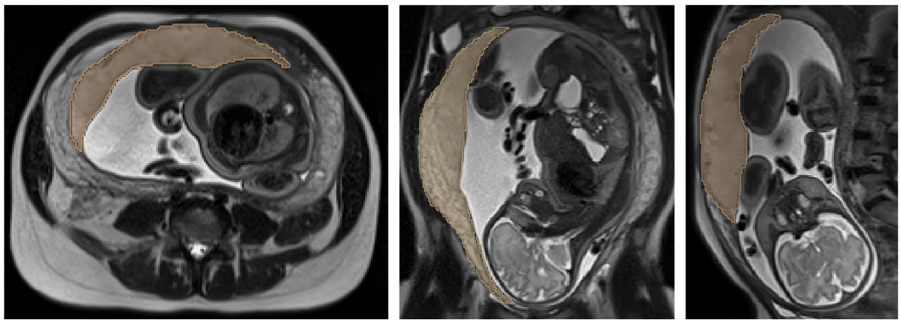

Fig. 4. Axial (*left*), Coronal (*middle*) and Sagittal (*right*) stacks of a fetal MRI. The placenta region is highlighted as a brown mask in each view.

pixel. Next, a Conditional Random Field (CRF) addressed the disconnectivity of labels resulting from the RF. With the scribbles being extended (around 250–300 mm), Slic-Seg rendered a stable and accurate segmentation with a DC = 0.79 ± 0.07, but it had a long runtime (42 s) because high-level features were used. The work published by Wang et al. (2016b) aimed at improving this methodology by combining multiple volumes acquired in different views. The authors refined the segmentation using the complementary resolution of volumes acquired from different views in a probability-based 4D Graph Cuts framework. A fast free-form deformation algorithm was used to register both sagittal and axial view volumes of the same patient. In order to overcome the inhomogeneous appearance between different slices and images, inter-slice and inter-image binary energies were defined based on the RF prediction. Therefore, Slic-Seg was more stable and less sensitive to the scribbles position. Although changes introduced increased the runtime (1.05 ± 0.13 s per slice) compared to Wang et al. (2015a) and Wang et al. (2015b), the performance was improved, yielding a DC of 0.82 ± 0.02. The intra-user *Fleiss' kappa* coefficient ranged from 0.931 to 0.949, whereas the inter-user *Fleiss' kappa* was 0.932, indicating high intra- and inter-user agreement with a low variability. Online RFs were also tested by Wang et al. (2016a) to make the aforementioned approach adaptive to imbalanced ratio changes. This combined a dynamically balanced online bagging to model the sequential arrival of data as a *Poisson* distribution, and a tree growing and shrinking strategy to update the online RF. The final probability of an unlabeled pixel was also combined with a CRF for spatial regularization, providing a DC = 0.89 ± 0.04.

The methodology discussed in these seminal works had several limitations. Firstly, although 4D Graph Cuts reduced the errors related to the initial propagation, the accumulated propagation error became higher in terminal slices due to a large shape change of the placenta between two slices. Secondly, the role of high-level features was to overcome the low contrast and weak boundary in the MRI data, which enabled a better description of the placenta and background than intensity-based low-level features. However, high-level information was insufficient to obtain a good segmentation and discrimination between the fetus and the mother tissues. Thirdly, to further correct the segmentations, user-feedback refinement should be considered. Another limitation is that there might remain a residual misalignment of the placenta between the different views. This could be improved by defining an inter-image neighborhood based on local voxels weighted by the distance or similarity. To overcome these flaws, the works published by Alansary et al. (2016) and Ceresa et al. (2017a) aim at avoiding the accumulated propagation error, applying learned high-level features to increase the robustness, being fully-automated, and exploiting the 3D neighborhood relationships. Alansary et al. (2016) proposed a fully automated segmentation framework with an extension to provide an intuitive pre-natal placenta view. Authors adopted a 3D multi-scale convolutional neural network (CNN) architecture consisting of eleven layers. Subsequently, a fully connected 3D dense CRF was employed to penalize inconsistencies of the previous segmentation. The first step in placenta segmentation recovery consisted in applying a flexible motion compensation algorithm based on patch-to-volume registration. The motion-free 3D image was then reconstructed using interactive super-resolution and 2D / 3D registration steps. Finally, a novel placenta visualization was supported by a mean-curvature flow skeleton and textured using techniques similar to curved planar reformation. A DC value of 0.72 ± 0.11 for 44 healthy fetuses, and 0.67 ± 0.15 for 22 healthy and IUGR cases were obtained. Notice that this multi-scale architecture provides larger 3D contextual information, which is essential for detecting highly variable organs. In this way, Miao et al. (2017) used the approach proposed by Alansary et al. (2016) to introduce an interior and exterior standardized visual representation of the fetal and maternal side of the placenta using a structure-aware slicing approach. To improve the segmentation, authors applied a post-processing based on morphological opening and closing operations to remove small local noise / objects and fill non-isolated holes.

Recently, a multiple-CNN approach called *DeepIGeoS* was designed by Wang et al. (2018b). They proposed to use two CNNs: an initial and automatic segmentation network and a refinement network. From the initial segmentation and some user interactions, the second network gives a refined segmentation. The interactions are encoded via geodesic distance maps and used as extra input channels of the CNNs. These distances help to differentiate neighboring pixels with different appearances, and improve label consistency in homogeneous regions. A resolution-preserving network was also presented to provide a better dense prediction by capturing high-level features from a large receptive field without loss of resolution throughout the CNN. The previous user interactions were finally integrated as hard constraints into back-propagatable CRFs, which applied freeform pairwise potentials rather than Gaussian functions. The proposed method led to accurate performances with fewer user interactions and less user time than traditional interactive segmentation approaches. Tested on 25 cases, it achieved a DC of 0.89 ± 0.05 and an average symmetric surface distance of 1.22 ± 0.55 mm.

Ceresa et al. (2017a,b) proposed an approach based on a 3D super-resolution of the uterus to automatically segment the placenta. Once the *Rician* noise and artifacts presented in the reconstructed MRI were eliminated, authors extracted robust texture features for the detection of the placenta using 3D Gabor filtering followed by GLCM. Finally, a binary SVM was applied to classify the placenta and non-placenta pixels. Related to this work, Torrents-Barrena et al. (2018) extracted the entire umbilical cord and provided an estimation of the cord insertion located on the placenta surface by means of a 3D recurrent fully-CNN. The method leveraged MRI inter-slice spatial dependences and convolutional

Fig. 5. An individual slice of a fetal 3D US. The placenta region is highlighted as a brown mask (*middle*), and with color Doppler displaying the blood flow (*right*).

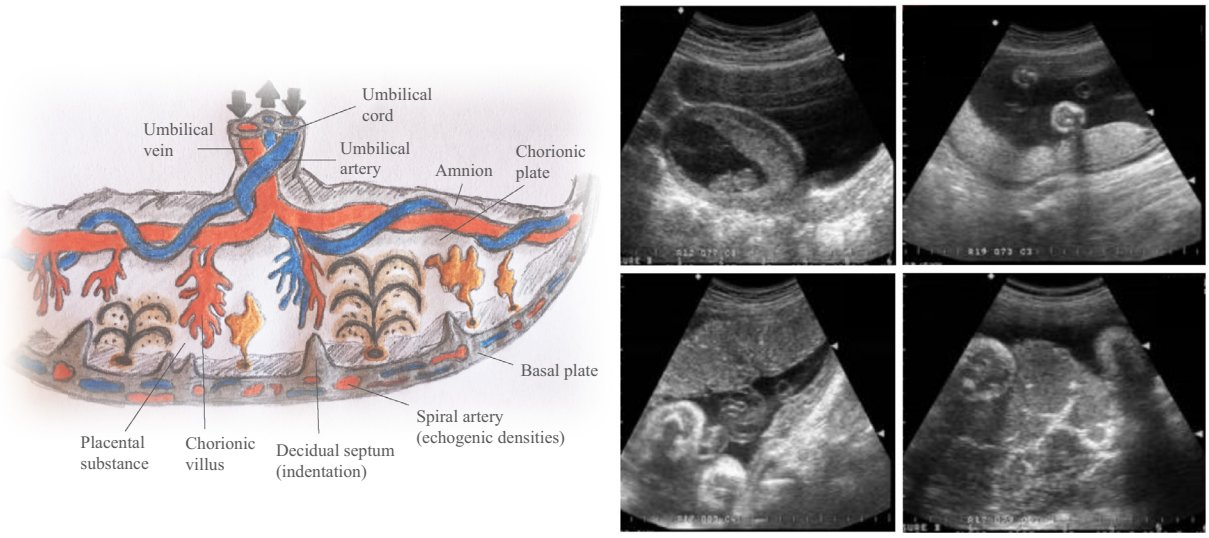

Fig. 6. Schematic view of the placenta (*left*) and visual perception of grades in *Grannum* classification (*right*): Grade 0 (*top-left*), Grade I (*top-right*), Grade II (*bottom-left*) and Grade III (*bottom-right*).

correlations through internal LSTM. Tested on 10 cases, it achieved a DC of 0.79 ± 0.03 and a Jaccard coefficient of 0.71 ± 0.02.

3.3.2. Placenta segmentation in fetal 3D US

Placenta segmentation through 3D US is nowadays performed clinically to explore if a significant relation between organ volume and fetal pathologies exists (see Fig. 5). In US imaging, placental shape can commonly appear irregular. In particular, in early pregnancies, where uterine tissue and placenta are very similar, the placental border measurement is challenging as only weak edges delineate the anatomical boundary (Roy-Lacroix et al. (2017b)).

Concerning 3D US, two works on semi-automated placenta volumetric segmentation were implemented by Stevenson et al. (2010) and Stevenson et al. (2015). The former (Stevenson et al. (2010)) does not provide details on the designed approach, but specifies that two different observers measured the placenta volume through various means: manually, using the virtual organ computer-aided analysis (VOCAL), and via their proposed volumetric method. Evaluation of inter / intra observer variability was performed. Authors provided concordant results for manual segmentation and outperformed VOCAL in terms of observer repeatability / agreement. The latter (Stevenson et al. (2015)) described the validation of a random walker algorithm. By weighting its graph edges with the intensity gradient, the stochastic movement of the random walker into heterogeneous intensity areas is avoided, thus favoring homogeneous regions of voxels. A first experiment revealed significant differences in the time taken to initialize VOCAL compared with the random walker method. A second experiment analyzed intra-/inter-observer variability using 3 observers. Results showed that intra-class correlation coefficient values for intra-observer variability were higher for the random walker method. The DC value for each observer was 0.84 ± 0.12, 0.94 ± 0.08 and 0.84 ± 0.11, respectively. However, the main limitations are: 1) only anterior placentas were studied; thus, the impact of anatomic variability in terms of placental positions cannot be evaluated. For a consistent placental volume calculation, the position of the organ is much more important than their appearance, and 2) intensive training is required to provide consistent results as Wang et al. (2015a), Wang et al. (2015b), Wang et al. (2016b) and Wang et al. (2016a). Oguz et al. (2016) presented a multi-atlas label fusion on 13 subjects. A rigid alignment of the atlases and target image using 3D US cone was performed after applying a median filter. This was followed by a bootstrapping approach based on a deformable registration between each atlas and the target. This procedure was repeated by using a ROI to maximize the registration ACC in and around the placenta. Finally, a post-processing eliminated geometric artifacts. Although a DC of 0.83 ± 0.05 was achieved, this approach was limited to anterior placentas as Stevenson et al. (2015).

Recently, Looney et al. (2017) demonstrated the feasibility of a CNN to segment automatically the placenta from 3D US volumes. They employed an open-source CNN called *DeepMedic* (Kamnitsas et al. (2017)). The ground truth data to train the network was derived from their semi-automated random walker approach (Stevenson et al. (2015)). Compared against their semi-automated segmentation, resultant median DC and Hausdorff distance were 0.73 (0.66, 0.76) and 27 mm (18 mm, 36 mm), respectively. Following a similar pipeline, Yang et al. (2017a) presented a threefold contribution. Firstly, they implemented a customized 3D fully CNN with dense voxel-wise semantic labeling to simultaneously segment the fetus, the gestational sac and the placenta. Secondly, due to the ubiquitous boundary uncertainty presented in

Table 1
Description of grades in *Grannum* classification.

Grade	Chorionic plate	Placental substance	Basal layer	Time of appearance
0	Straight and well-defined	Homogeneous	No densities	Early 1st trimester to 2nd trimester
I	Subtle undulations	Few scattered echogenic areas (EGAs)	No densities	Mid early 2nd trimester to 3rd trimester (18 - 29 gestational weeks)
II	Indentations into placenta	Linear echogenic densities	Linear arrangement of small EGAs	Late 3rd trimester (30 gestational weeks)
III	Indentations communicating with basal layer	Circular densities with echo-spared areas in center; large irregular densities, which cast acoustic shadows	Large + confluent basal EGAs with acoustic shadows	30 gestational weeks post-dates

US, a bidirectional long short-term memory network was designed to sequentially refine the corrupted intermediate probability volumes provided by the previous CNN. Thirdly, a hierarchical deep supervision procedure was introduced to adequately boost the data flow within the recurrent neural network. A DC of 0.643 was reported for placenta segmentation. In a more recent work, a fully CNN named *OxNNet* was trained using 2393 first-trimester placentas (Looney et al. (2018)). This exceptionally ground-truth was segmented through a random walker. The *OxNNet* was inspired by a U-Net / V-Net architecture. The number of convolutional layers and channels utilized was customized and max-pooling was replaced with stride convolutions. The DC monotonically increased from 0.73 for 100 training samples to 0.81 for 1200 cases.

3.3.3. Placenta classification in fetal US

Gestational diabetes mellitus (GDM) appears when women without diabetes, develop high blood sugar levels during pregnancy. Usually, few symptoms emerge, but it rises the pre-eclampsia risk, depression, and the chance to need a Caesarean section. Babies born from these mothers present a high risk of being too large, having low blood sugar, and jaundice. In GDM, the placenta suffers several structural and functional alterations such as an increment of its weight, central thickness and diameter. In the past, these alterations were simply inspected visually by an experienced doctor. To prevent visual fatigue and reduce errors, *Grannum's* classification (Grannum et al. (1979)) was designed to categorize the placenta changes in GDM and high-risk pregnancies.

Hence, placental grading (Grannum et al. (1979)) refers to a US-based four-stage system based on placenta maturity (see Fig. 6). This primarily affects the extent of calcifications and fibrosis degree. Each grade is defined in Table 1.

The classification of the left and right US placenta images into normal and GDM was performed by Malathi and Shanthi (2009) and Malathi and Shanthi (2011) using statistical measurements (*e.g.,* area and perimeter of the segmented placenta, mean, contrast, correlation, entropy and sum of squares). A feature dimensionality reduction was carried out through a Haar wavelet decomposition of 5 levels. A watershed approach and a neural network were then applied to obtain the final segmentation. Similar statistical features were also extracted and classified in Malathi and Shanthi (2010c) considering that maternal GDM is associated with an increment of placental thickness. The gestational age was estimated using crown-rump length and biparietal diameter in the first trimester, biparietal diameter and femur length in the second, and biparietal diameter, femur length and abdominal circumference in the third trimester. Correlation and regression analysis were used to establish the relationship between placenta thickness and gestational age. A remarkable increase in weight, diameter and ventral thickness was found in placenta of diabetic mothers. In addition, authors found that placenta had a roughly oval or rounded shape in all groups except in those suffering from GDM complications.

Other authors went deeper in the task of classifying placenta US images according to the 4-grading proposed by *Grannum*. Linares et al. (2003) used the ability of a decision tree classifier to discriminate different textural features such as co-occurrence matrices, Law's operators, and neighborhood gray-tone difference matrices (NGTDM). Results showed an ACC of 87%. Similarly, in Linares et al. (2004) the same set of features was computed and weighted separately using the relief-F method. The classification strategy adopted was the k-Nearest Neighbor (KNN). An ACC of 60.71% was obtained using Law's and NGTDM features. Instead of extracting NGTDM features, the work proposed by Chen et al. (2011) utilized 72 texture features derived from the spatial gray-level dependence and difference matrices to determine quantitatively the gestational age. In addition, they incorporated the contrast, angular second moment, entropy, mean, and inverse difference moment. To assess the relationship between the aforementioned features and the gestational age, a multiple linear regression model was applied. Histogram intensity level and textural features such as mean, standard deviation, contrast, correlation and entropy were also explored in Malathi and Shanthi (2010a) and Malathi and Shanthi (2010b). Since inner and boundary pixels stand for inner texture regions, segmentation was achieved by applying an adaptive region growing. The performance of grayscale features such as mean, variance, distortion and kurtosis was also investigated by Liu et al. (2009). A multi-class SVM based on a radial basis function was fed with the previous features. Authors achieved a total recognition rate of 92%. They also realized that *grade III* was often the most confusing stage due to its subjective visual characteristics.

Efforts to improve the recognition rate of *grade III* were made by several researchers. They discovered that calcifications were predominant in placentas belonging to *Grannum grades II* and *III*. Early preterm calcification was therefore shown to be associated with a higher incidence of adverse maternal and fetal outcomes. Ryan et al. (2008) developed a software tool to analyze the placenta in search of calcifications. Firstly, a flood-filling algorithm was applied to create a reference image, where saturated pixels were turned black and the others, white. In order to boost the classification of placental borders, the *Bresenham*'s line drawing algorithm was utilized. A region-growing was then applied to find calcifications against a pre-studied threshold. Furthermore, the selection of individual placental infarcts was enabled through the implementation of a circular region of interest (ROI) tool. The same objective was shared by Moran et al. (2013), in which a manual placenta delineation was performed to identify calcifications through a user-controlled intensity threshold. The doctors' reproducibility achieved a DC of 0.89 $\pm$ 0.06 (involving all permutations for all observed-pairs). Afterwards, a flood-filing algorithm was employed to create a secondary reference map to be used during the quantification. An area under the curve equal to 0.628 was finally obtained for placental grading.

The lack of objectivity, precision and reproducibility of the *Grannum*'s method (Moran et al. (2009)) provides a vague definition of *grades II* and *III*, which considerably affects their final classification accuracy. Although some aforementioned works (Liu et al. (2009)) presented good recognition results, the features do not exactly describe each stage individually due to the qualitative definition of this categorization. To overcome this issue, authors propose an index based on the selected features as a quantitative alternative to *Grannum* classification. To this end, Li et al. (2014) proposed another method to automatically stage the placenta maturity. Dense sampling was applied to pre-process all images, and after patch extraction, each feature was calculated through the local region invariant DAISY descriptor. Features were encoded to reduce the dimensionality. Afterwards, these features were fitted to a Gaussian mixture model and clustered via K-means. These clusters were then fed to a SVM for the final stage decision. An ACC of 0.874 was obtained, as well as a SEN = 0.996, and a SPE = 0.874. Similar to the previous article, the method developed in Lei et al. (2014) was also aimed at extracting local intensity order pattern features in each image region. The Harris-Laplace algorithm was then adopted to detect placental key points. This time, features were encoded using a Fisher vector. The input placental image was partitioned into patches, and each patch was represented by a local intensity order pattern. The rest of the methodology was exactly the same as in Li et al. (2014). Final results presented an improved ACC of 0.9375. In this case, authors achieved a lower SEN = 0.9804, but a higher SPE = 0.9375. Finally, the article in Lei et al. (2015b) summarized the results of Li et al. (2014) and Lei et al. (2014). Several future research lines were suggested such as the use of blood flow information to improve the whole placenta evaluation, and the fusion of classifiers to accurately recognize all placental grades.

3.4. Heart

The finding of a heart defect in the fetus has significant implications for pregnancy management, delivery planning, and diagnosis of abnormalities in other organs. Congenital heart disease (CHD) means a heart condition or defect that develops in the womb, before a baby is born. There are many different types of CHD such as aortic stenosis, atrial septal defect, hypoplastic left heart, patent ductus arteriosus, supraventricular tachycardia, among others (Hunter and Simpson (2014)). Some CHD problems are currently detected when the mother has an US scan during pregnancy (usually at 20 weeks of gestation), but sometimes they are not discovered until the birth, or even until the child becomes adult. In babies and toddlers, CHD can have a range of symptoms (extreme tiredness, poor feeding, excessive sweating, rapid heartbeat, breathing problems, chest pain and a blue tinge to the skin) because every child and condition is different (Postma et al. (2016)). Treatment as medication or fetal heart surgery depends on the type and severity of the condition.

An accurate fetal heart segmentation is essential to early locate such structural heart abnormalities before birth. This is difficult due to the small size of the fetal heart structure and cavities, especially the thin chamber boundaries in the atrial septum, the membranous segment of the ventricular septum, and the valvular leaflets.

3.4.1. Heart segmentation in fetal MRI

Fetal cardiac MRI can provide additional information when compared to echocardiography. In particular, MRI offers a better volume estimation of the heart chamber, since US measurements are often inaccurate due to acoustic shadows. Additionally, the MRI four-chamber view is widely used as a main screening method for fetal cardiac defects because of its large field of view. The segmentation of the fetal heart in MRI is currently a cutting-edge research topic, and only one work has been published so far. Once it is detected, an accurate 3D volume can be rendered to clinically examine any abnormality in the organ.

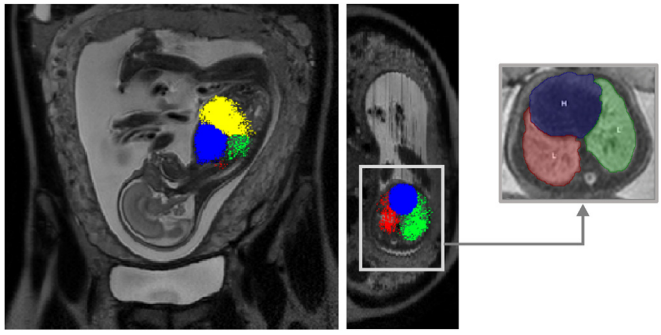

Fig. 7. Figure adapted from Keraudren et al. (2015): Rough organ segmentation of a fetal MRI: heart (*blue*), left lung (*red*), right lung (*green*), liver (*yellow*). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Keraudren et al. (2015) proposed to automatically detect the heart and also the lungs and the liver using random forest and steerable features (see Fig. 7). The main challenge for the fully-automated localization of them was the unpredictable orientation of the fetus. All MRI were mapped into a standard coordinate system defined by landmarks on the fetal anatomy and normalized for fetal age. Assuming that the brain center was known, image features were extracted in a coordinate system rotating around the brain to find the heart. The lungs and the liver were then localized using a coordinate system in which one axis was parallel to the heart-brain axis. Authors validated their method on both healthy and IUGR affected fetuses. The detection rate was beyond 90% for all organs belonging to healthy fetuses provided that motion artifacts were not present. The detection rate decreased to 83% for the heart, 78% for the lungs and 67% for the liver when motion was present. The method could be used to initialize subsequent processing steps such as segmentation or motion correction.

3.4.2. Heart segmentation in fetal US

For first-trimester fetus, US imaging is limited by the small heart size and the low signal-to-noise ratio. Due to these limitations, the whole fetal cardiac structure detection (see Fig. 8) is complex even for experienced doctors. Furthermore, fetal ribs shadowing causes heavy dropouts behind the bony structure. The four-chambers volume quantification is therefore complicated by these artifacts that difficult the determination of cardiac indices (i.e., ejection fraction and stroke volume). Many automatic segmentation algorithms can leak through these missing boundaries if intra-chamber walls are unresolved.

Dewi et al. (2014) reviewed of 2D fetal echocardiography techniques. The authors focused on three main topics to support diagnosis: speckle reduction, image segmentation, and image analysis. The section related to image segmentation included algorithms such as active contours, level sets and deformable models for heart tracking. They also remarked the importance to measure fetal cardiac characteristics (heart position, cardiac axis, situs and structural delineation) in abnormal diseases such as CHD. In this regard, segmentation algorithms can aid the quantification of geometrical parameters like area, volume and motion cardiac structure. Dewi et al. (2014) reviewed 14 segmentation works, which are therefore not discussed here.

Among the pioneering works that have addressed the segmentation of fetal echocardiographic images, Piccoli et al. (1999) discussed several supervised and unsupervised approaches. On

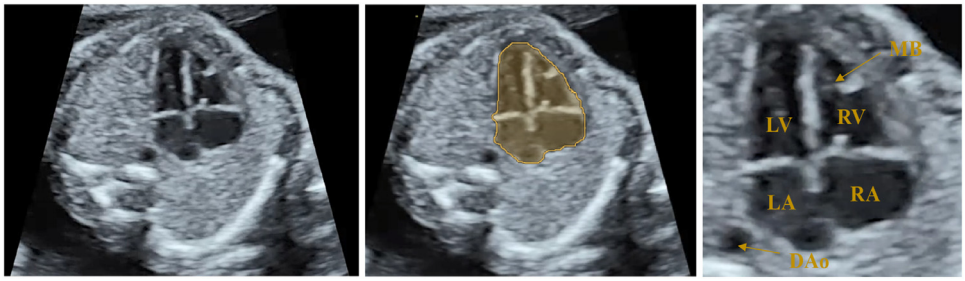

Fig. 8. Fetal echocardiography: original four-chamber view (*left*), heart region masked in yellow (*middle*), and heart parts (*right*): left ventricle (LV), right ventricle (RV), left atrium (LA), right atrium (RA), moderator band (MB), and descending aorta (DAo). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the one hand, they applied a back-propagation neural network, which received different textural features (mean, variance, contrast, entropy and homogeneity) as input. On the other hand, they proposed an unsupervised network based on self-organizing maps (SOM) and the K-means technique. Similarly, LemosDeSiqueira et al. (2002) focused on the use of the unsupervised SOM to approximate the probability density function of the US patterns. The map was also post-processed through the K-means clustering. Each segmented US image was generated by correlating its pixels and clusters found in the map. This work differs from the previous (Piccoli et al. (1999)) as the authors only chose mean and variance to train the neural network. To refine these measurements, the least-means squares error selected the curve that best fitted the pixels located at the border of the ventricle wall. Additionally, the authors developed an algorithm to measure semi-automatically the width of the intraventricular septum and the left ventricle cavity. Results were affected by ventricle internal wall discontinuities called trabeculations. Although the supervised methods identified some cavity and tissue regions, specially the inter-ventricular septum, important information such as regions near to the class / tissue border was lost. These deficiencies were probably caused by incorrectly selected tissue and cavity samples in the noisy US, or inherent information loss when complex US were segmented in just 2 classes. Thus, US segmented into more than 2 classes by unsupervised approaches tend to better preserve heart structures. Subramanian et al. (1997) combined the region growing and the split / merge technique to segment 6 US by speeding up the segmentation via frame-to-frame video coherence. However, 4–18% of images were wrong.

Differently, Tutschek and Sahn (2008) aimed to search a novel obstetric application of their method to semi-automatically segment and visualize fetal heart cavities (the ventricles, atria and great arteries) and vessels. The volumes were studied offline using a commercially available product designed for *ovarian folliculometry*. During the analysis, individual cavities in a static volume were selected. Once the automatic segmentation was done, the system started a search to isolate anechoic or hypo-echoic spaces. Then, self-expanding volumetric seeds were placed automatically. The main limitation in the method is related to acoustic shadowing, which produced hypo-echoic segments and may erroneously indicate additional volumes. In addition, it lacked a fully-manual mode to either place or edit volumes. Due to operator-dependency required by manual editing, this approach is not suited for routine fetal echocardiography.

The active contour algorithm was also employed in echocardiographic images by Lassige et al. (2000) to automatically extract the endocardial borders of septal defects. Furthermore, the size of the defects in the original 2D and 3D datasets were also compared. The methodology required the selection of 2 points that had to accomplish the following criteria: 1) one point must be located on each side of the septum, 2) the points should not be placed against a potential border of the defect, 3) the points should not be placed excessively far from the septal defect, and 4) the points should be located not just in each of the ventricles between which the septum was located, but the defect itself should be located on a line that connects both points. Results were valid for approximately 70–85% of the tests. It was noted that 3D US seemed to underestimate the defects size and may fail to detect small defects.

Most published works, however, employed level set approaches instead of active contours to segment the chambers of the fetal heart. Dindoyal et al. (2007b) evaluated the performance of 3 existing level set methods: 1) edge-based, 2) region-based, and 3) shape-based level sets. As the main advantage, shape-based level sets together with shape prior constrained the propagation to a predefined shape in areas of ambiguity such as missing chamber walls or spurious edges. When the prior was turned off, shape-based level sets segmented the cardiac boundaries, but the weak walls between chambers were punched due to the lack of a collision term. Unlike region-based level sets, this algorithm did not require any coupling between the individual snakes, and it minimized the inter-chamber leakage. The region-based level sets appeared to perform similarly with and without the shape prior term as its collision detection abilities handled the weak boundaries in-between cardiac chambers. Without shape prior the edge-based level sets completely overshot the boundaries due to its constant advection term. Hence, it was concluded that the snake algorithm with a shape prior term is more robust across many images compared to the amorphous snakes. In each snake type, the shape prior term yielded a better segmentation agreement being the greatest improvement in the edge-based level sets. Among their findings, the same authors outlined an implicit level set deformable model with shape prior term (Dindoyal et al. (2007a)). Afterwards, a *Mumford-Shah* model was combined to minimize the region-based energy function. All segmentations were within 1.5 - 2 pixels of error compared to manually traced contours. In the case of no boundary due to unresolved structures, the shape prior was able to provide an estimation based on stored shape knowledge. Sriraam et al. (2012a) also suggested the application of amorphous snakes, in which the shape prior term improved the individual identification of the chamber edges and contours. Another approach that inspected an unconstrained level set deformable model was designed by Dindoyal et al. (2011). It was penalized from growing into an adjacent cardiac compartment through a novel collision detection term. An advantage over the statistical model was that the presented method can segment a high variation of developing hearts that were not considered in the training set. Hence, errors appeared larger in 3D than for 2D by a discrepancy of 0.5 mm. The left and right ventricles segmentation for the computation of ejection fraction provided a 95% of agreement with the clinical tracings. In the case of fetal heart US images, the left ventricle usually showed the best overlap due to its size, elongated shape and less wall dropout. In the right ventricle, the papillary

muscle made the manual segmentation more challenging. The atria was even more difficult because of the septum.

In the extraction of cardiac structure patterns is important the selection of a suitable ROI. Systems developed by LemosDeSiqueira et al. (2007) and Deng et al. (2010) firstly computed an entropy texture feature to decide and isolate the most convenient ROI. Once the candidate regions were found, the position of the pattern with a higher degree of similarity with respect to a fetal heart was kept. The last stage was based on the search of the ROI inside the US image. To locate the cardiac structure, a comparison between the previous candidate mold and the heart structure mold generated in the first stage was performed through the *Bhattacharyya* coefficient. The authors were satisfied with a significant ACC of 89%. Deng et al. (2012) proposed an extension of their previous work (Deng et al. (2010)). The improved approach consisted of 2 main parts: 1) pre-processing phase: a) a rectangular mask was used to find the location with the highest level of accumulated motion based on a Gaussian pyramid, and b) through the 3D combination of *Rayleigh-trimmed* filter and anisotropic diffusion, a despeckling method was developed to eliminate the speckle noise and emphasize the motion information, 2) the active cardiac model was designed to detect the fetal heart. The key contribution of this article was that both the structure and the motion information were considered simultaneously. Authors assumed that a fetal heart consists of eight parts: left septum, upper left boundary, top boundary, inter-ventricular septum, upper right boundary, right septum, lower right boundary, and lower left boundary. To represent the spatial relationship among different parts, a center part was added and connected with the other 8 parts. The connections between parts at the same time represent the structure information and the connections between parts at different times represent the motion information. The extended method had an error lower than 13 pixels. Based on the motion information, Vijayalakshmi et al. (2013) superimposed consecutive frames to design the region mask of the four heart chambers. The dissimilarity between frames (different phases in a cardiac cycle) was indicated by comparing the amount of residual gray shades distribution of the chamber region. Once the superimposed image was converted into a binary image, connected component labelling was applied. Using the information of the largest connected component, all edges of the heart chamber region were identified. With the same purpose, Sriraam et al. (2012b) and Reshmi-Mariam-Reji-Jacob et al. (2013) involved an automatic ROI localization after a noise removal step based on morphological operations and the median filter, respectively. The latter also employed an effective K-means clustering and an active appearance model to finally identify the fetal heart. Therefore, small ROIs did not represent adequately the fetal heart structure due to the present noise, the variability of the heart size, and the fact that the search was exclusively based on gray intensity. Furthermore, the performance achieved by Deng et al. (2012) was better than that obtained by the active appearance model alone. The reason was that the latter did not consider the motion information for each part and was mainly based on the appearance information. When the data were collected at later gestational ages, the ACC increased, which was reasonable considering that the cardiac structure becomes bigger and better defined in older fetuses.

Several authors invested their efforts on modeling the shape of the left ventricle. Vargas-Quintero et al. (2016a) designed a multi-texture active appearance model based on the steered Hermite transform to efficiently code texture patterns of the input US scans. Afterwards, different methods such as thresholding, mathematical morphology and correlation were applied to localize the left ventricle. Results reported a DC of 0.8631 and a point-to-curve distance of 2.027 pixels. Similarly, the Hermite transform was also employed by Vargas-Quintero et al. (2016b) to minimize the speckle noise present on US images. A point distribution model was then trained to detect the fetal left ventricle. Likewise, Vargas-Quintero and Escalante-Ramírez (2013) compared the previous Hermite transform denoising with a wavelet decomposition approach as a way to strengthen the performance of the fetal left ventricle segmentation. For the latter, a Bayesian estimator at subband level for pixel classification was employed. Next, authors picked out a deformable model to evaluate the influence of the aforementioned filtering techniques.

Recently, Yu et al. (2017) compared the use of active and dynamical appearance models with dynamic CNNs to provide a robust segmentation of the left ventricle. The first US frame was delineated manually to fine-tune the CNN, and the other frames applied a shallow tuning to adapt to the fetus. To differentiate the left ventricle and atrium regions, the mitral valve base points were tracked. Experimental results showed a good segmentation agreement of DC = 0.94 with the ground truth, especially in images with leakage, blurry boundaries, and subject-to-subject variations. Following the pipeline illustrated in the previous work, fully CNNs were also used by Sundaresan et al. (2017) to detect the fetal heart and categorize each US frame as belonging to the standard viewing planes. The proposed CNN relied on a deep 16 layer VGG architecture. To compare the final error rate of several fully CNN models, predictions were down-sampled by a factor of 32, 16 and 8 to obtain pixel-wise segmentation. In the training stage, the parameters of each layer were updated using stochastic gradient descent and softmax log loss function. Afterwards, a two-stage post-processing was designed to eliminate spurious labels: 1) multiple non-background labels, which were close (20 pixels) to an image edge, were removed, and 2) labels smaller than 1200 pixels were ignored. Their model obtained a classification error rate of 23.48%. Another deep learning approach was proposed by Huang et al. (2017), which presented a three-fold contribution: 1) a CNN was developed to predict the visibility, viewing plane, location and orientation of the fetal heart at the frame level, 2) an anchor mechanism with an intersection over union loss were both implemented to increase the localization accuracy, and 3) a bi-directional LSTM with 256 dimensional hidden layers was designed to temporally extract convolutional features, allowing each prediction to be conditioned on the entire fetal cardiac US video. The model gave an error of 27.7%. Gao and Alison-Noble (2017) also designed an end-to-end two-stream fully CNN for temporal sequence learning to recognize, characterize and fuse spatio-temporal fetal heart representations. The network captured motion and appearance features in a weakly supervised manner. In terms of heart identification, the obtained results for 412 US videos demonstrated a 0.90, 0.85 and 0.89 of ACC, precision and recall, respectively.

Sampath and Sivaraj (2014) involved a probabilistic patch-based maximum likelihood estimation denoising approach as a pre-processing step to eliminate the inherent speckle noise. Then, a fuzzy connectedness algorithm was applied with the use of user-predefined seed points inside the fetal heart. Statistical measures such as mean and standard deviation of the seeds were calculated to extract the features. The segmentation results provided a fuzzy connectedness map defined by the pixels strength of hanging togetherness (*spels*), where the object of interest was extracted through an image thresholding. Hence, boundaries and edges were characterized by fuzzy *spels* of heterogeneity nature, which remained unconnected with their adjacent fuzzy *spels*. The strength of fuzzy connectedness in contiguous regions between two *spels* was described as the fuzzy *spel* affinity using a combination of 3 parameters: 1) The unit of coordinate space adjacency, 2) The unit of intensity space adjacency, 3) The unit of gradient intensity space adjacency. The average DC after the probabilistic patch-based maximum likelihood estimation was equal to 0.9795, and without it the DC was 0.186. Indeed, in recent years, more sophisticated

techniques appeared to deal with specific heart diseases. For example, fetus cardiomegaly, an abnormal enlargement of heart, is one of the early indicators of hemoglobin *Bart*'s disease. To solve this non-previously tackled issue, Sardsud et al. (2015) utilized the *Horn-Schunck*'s optical flow algorithm to extract a first ROI (that was supposed to contain the heart chambers) to find the frames with end diastolic and end systolic. Then, a second ROI, supposed to have the fetus chest, was also extracted. They also performed patch-based C-means with 20 clusters in the second ROI. Similar patches were finally combined together to segment the heart chambers. To reduce some small areas of noise, an opening filter was applied and then, objects that were connected to the boundary of the second ROI and did not overlap with the first ROI were also discarded. The authors found out a segmentation error around 2.17% compared with an expert's opinion. Some error occurred when the second ROI covered a larger area than expected, since the algorithm selected the wrong structuring element size in the opening filter. In this case, the algorithm discarded some correct areas, producing smaller heart chambers.

Sridevia and Nirmala (2016) developed a framework to diagnose the prenatal truncus arteriosus congenital heart defect for the first time. The system started with the suppression of undesired distortions through the probabilistic patch-based maximum likelihood estimation. Afterwards, anatomical structures were highlighted using the fuzzy connectedness segmentation process. This delineated blood vessels starting from a single seed point / *spel* provided by the user. A total of 32 diagnostic features were extracted: statistical and Haralick features, laws textural energy measures features, neighborhood gray tone difference matrix features, Fourier power spectrum features, and shape parameter features. A subset was then selected via Fisher Discriminant ratio analysis. To reduce the error rate, authors firstly chose the best 5 features to define the most optimal subset. Finally, a new adaptive neuro-fuzzy inference method was built to classify the previous top ranked features as normal / abnormal fetal heart. The final average error rate was: an area under ROC = 0.8954, a F-score = 0.9673 and a Kappa = 0.7914. As the main drawback, the proposed inference system was specifically built to diagnose the presence or absence of prenatal truncus arteriosus congenital heart defect among 13 various types of CHD. In addition, the work published by Nirmala and Sridevi (2016) determined the H-shaped sonographic marker to recognize the prenatal ventricular septal defect. The methodology involved cropping out the ROI followed by a probabilistic patch-based maximum likelihood estimation to eliminate the noise. Then, an unsupervised Markov random field approach was utilized to detect the fetal heart chambers. It yielded an error percentage of 9.83%. Bridge et al. (2017) also presented a method for the antenatal diagnosis of CHD. To overcome the ambiguity of appearance of fetal structures in US images, they used a sequential Bayesian filtering solution. Intensity features provided a fast but less accurate prediction of the fetal heart orientation and the cardiac phase (e.g. classification error of around 50%), whereas the use of gradient features presented a slower runtime and a lower error rate (e.g. classification error of around 30%). The inclusion of motion features reduced the classification error to under 20% and improved the cardiac phase prediction, but with a significant reduction in speed. Concretely, intensity features performed better when more features were available. By contrast, gradient features gave rich description with a reduced number of features, but implied a large reduction in speed. Classification and regression forests were trained with the aforementioned features to predict the visibility, location and orientation of the heart. A modification of such regression forests for circular variables was implemented to deal with the cardiac phase prediction. Maraci et al. (2017) also employed the method presented by Bridge et al. (2017) to automatically detect the fetal heart on 323 free-hand US with a mean ACC of 72.92%. In order to identify the heartbeat, a frame by frame categorization followed by a CRF regularization was conducted. A subsequent kernelized linear dynamical model achieved a mean heartbeat classification ACC of 93.1%.

3.5. Lungs

Congenital lung disorders, also known as cystic lung disease or congenital lung malformations, occur while the baby is still in its mother's womb. Lungs are almost the last organs to be developed during the prenatal stage. Concretely, they are not fully-grown until the end of pregnancy. A deeper understanding of the pulmonary anatomy in the premature neonate is achieved by conducting a longitudinal study of normal lung maturity. Fetal lung maturity is determined mainly by pulmonary surfactant and it can only be assessed with laboratory tests on amniotic fluid (DeSilva et al. (2016)). The need of non-invasive techniques to robustly predict the lung maturity has recently emerged.

With the increased use of obstetric US and MRI scans, congenital lung disorders are detected more often antenatally, which allows for proper planning of peripartum and neonatal management (Bhattacharya and Mariani (2013)). Prenatal treatment options include the maternal administration of steroids, minimally invasive procedures or open fetal surgery. These interventions aim to alleviate the mass effect, prevent the progression of complications and improve the outcome for affected fetuses.

3.5.1. Lungs segmentation in fetal MRI

The first semi-automatic method for fetal thorax analysis in genuine MRI volumes was developed by Kainz et al. (2014b). A fast rotation invariant spherical harmonic descriptor was designed to train a forest ensemble method for the spina cord localization. As the heart and spinal cord positions were known, a probability map was generated for lung voxels based on a graphical model of the thoracic organs. Authors used logistic regression to predict the gestational age specific distances between the heart and the left / right lung. Subsequently, voxels that belonged to each class as initial foreground / background constraints were defined to apply a geodesic active contours segmentation. Spinal cord and lung segmentation achieved a DC of 0.89 and 0.87, respectively. An active contour approach was also employed in Chikop et al. (2014) to render a lung volume for illustrating biometric measures required to determine fetal developmental rates in pathology characterization. Segmentation was followed by morphological operations, which removed spurious regions. The banding artifact was eliminated through Haar wavelet and Fourier transform filtering. In some cases, these methods obtained a truncated lung extraction, as certain parts were not visible in the T_2 weighted scans due to the early gestational week. Since the segmentation did not rely on a-priori information of the image based on hyper or hypo-intensities, the proposed approach would need to be extended to deal with T_1 weighted data. In addition, authors realized that when lung volumetry was performed on motion corrected high resolution reconstructions, the performance appeared to be closer to post mortem measurements.

Rajchl et al. (2017) also followed the idea of using an energy minimization function such as the one applied in the active contour model to create a modified graph cut. They extended its concept with recent advances in CNN modeling and proposed *DeepCut*, an architecture to recover semantic segmentations of images with corresponding bounding boxes. Authors designed a simple CNN inspired by *LeNet* architecture. Because of anticipated motion artifacts between fetal MRI slices, they decided to emphasize the in-plane context of their training patches. Post-processing included regularization with a densely-connected CRF. The two proposed *DeepCut* initialization methodologies achieved higher

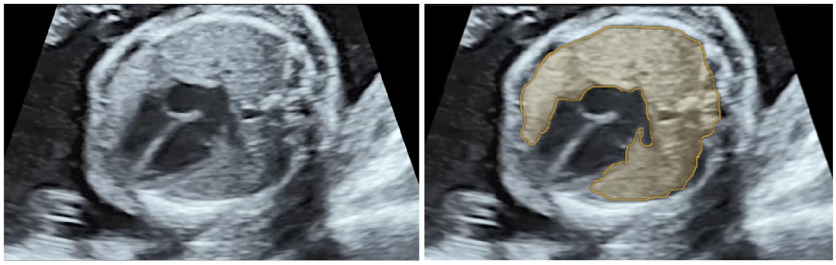

Fig. 9. An individual slice of a fetal chest US (*left*), the lungs region is highlighted as a brown mask by an expert clinician (*right*).

average ACC (70.0 ± 8.1% with bounding boxes, and 74.9 ± 6.7% with pre-segmentation), however the ACC remains lower than fully-supervised CNN (82.9 ± 10.0%). By restricting learning background patches, the architecture avoids learning features for the entire image domain, allowing for faster training. Expressing the employed CRF as a recurrent neural network might be a solution for back-to-back training of the parameters involved in each *DeepCut* iteration. However, the iterative nature of the previous algorithms was limited to find local optima. Thus, the resulting segmentation was dependent on the initial regions. A deep-learning framework was also implemented in Wang et al. (2018a) to provide a 2D segmentation of the fetal lungs, placenta and brain. Authors coupled CNNs incorporated into a bounding box and a scribble-based pipeline. The proposed method outperformed traditional interactive segmentation approaches thanks to a weighted loss and an image-specific fine-tuning which, in turn, make the CNN model adaptive to any test image. Besides the improved ACC (achieving a supervised and unsupervised DC of 0.92 ± 0.03 and 0.85 ± 0.06, 0.92 ± 0.03 and 0.86 ± 0.08, and 0.96 ± 0.02 and 0.90 ± 0.06 for lungs, placenta and brain, respectively), fewer user interactions and shorter computational times were needed.

As we explained in Section 3.4.1, the method proposed by Keraudren et al. (2015) was also aimed to segment the fetal lungs.

3.5.2. Lungs segmentation in fetal US

Segmentation of fetal lung in US images (see Fig. 9) has become an ambitious task due to low signal-to-noise ratio of this modality. In particular, the speckle noise and the signal dropout affect fetal chest US, which results in a deficient contrast and blurry boundaries (poor image quality). A suitable fetal lung segmentation is also hard to achieve because of the similar tissue densities that the lung and heart myocardium have. Nevertheless, researchers are more focused on finding the interrelationship between gestational age and its corresponding normal fetal lung volume.

There is only one method, developed by Li et al. (2010), to semi-automatically segment the fetal lung from chest US images. The procedure was firstly proposed to enhance boundary regions through multiple GLCM textural features such as contrast, correlation, entropy and homogeneity. The median and variance were also calculated as descriptors of the region. Pixels were classified into 3 categories: 1) background (with low intensities), 2) strong boundary regions (with high intensities, corresponding to real tissue boundaries), and 3) weak boundary regions (with moderate intensities corresponding to tissue details). For this purpose, a Gaussian mixture model of 3 independent Gaussian distributions was adopted. The boundary skeleton was calculated through a morphological thinning process. Since the segmentation was a multi-object detection task, to avoid interferences, 3 rectangles were manually defined as ROIs for the fetal chest, the heart and the spine. These served as initial contours for the vector field convolutional model. Thus, the fetal lung was the region within the fetal chest, but excluding the fetal heart and the fetal spine. Further work is needed to properly validate this approach, as the results presented were qualitative and no performance measures were calculated.

3.5.3. Lungs classification in fetal US

Several studies have demonstrated that texture analysis of fetal lung US images based on gray-level measurements, lung tissue motion and relative features of lung-to-placenta or -liver images are able to identify feature patterns that correlate strongly with gestational age, or with the results of fetal lung maturity tests on amniotic fluid.

Bhanu-Prakash et al. (2002) described the first general attempt to analyze the fetal lung maturity using US images to obtain the textural trend of normal fetal lung as a function of the gestational age. Regions of interest from lung and liver were utilized to extract several texture features such as fractal dimension, lacunarity from fractal measures, mean, standard deviation, coefficient of variance, and second moment calculated from the histogram. The ratio of these features was used as input to several well-known classifiers: nearest neighbor, KNN, modified KNN, multilayer perceptron, radial basis function network, and SVM. An ACC from 73% to 96% was obtained. Although classifiers were not biased by data, the neural network and its variants showed a poor generalization capability. The performance was better in the case of radial basis function network and SVM classifiers. When the training and test data contained samples from all gestational ages, the neighborhood classifiers performed better due to more closeness in the neighborhood relation among the samples at the boundary of the classes. Along the same lines, Cobo et al. (2012) explored the feasibility and reproducibility of the features extracted from their AQUA software to provide gestational age information. This algorithm was also based on a texture analysis, which obtained 15.300 features per delineation, and a CRF. A statistical learning algorithm was then applied to select only the most 30 relevant features. Once these features were chosen, a regression model composed by random forests was computed to test the features' robustness. The association between gestational weeks and AQUA lung textures yielded a *Pearson* correlation of 0.97. Palacio et al. (2012) implemented a second release of the AQUA software built upon genetic algorithms and SVM. The former identified a set of 31 critical descriptors, and the latter developed a mathematical model using these descriptors to predict the lung maturity in amniotic fluid. Average ACC of 90.27% was obtained, as well as a SEN and SPE of 95.1% and 85.7%, respectively.

AQUA software therefore provided information about lung structural changes in transverse US without any reference tissue to compute relative echo-densities. No blind samples were used to test the predictive capacity of the model. Although the previous study provided a proof of principle that fetal lung maturity assessment can be achieved through a non-invasive method, the risk of respiratory morbidity was not considered despite its clinical importance. Recently, Bonet-Carne et al. (2015) tackled this issue using neonatal respiratory morbidity cases, in which some corresponded to respiratory distress syndrome and others to transient tachypnea of the newborn. A pre-defined method termed

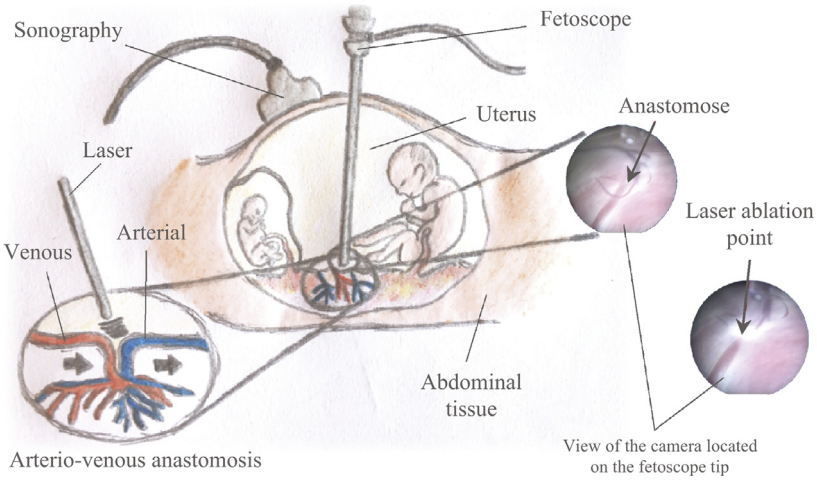

Fig. 10. TTTS fetoscopic laser photocoagulation. This US-guided therapy consists in closing blood vessels (anastomoses) connecting the twins. The selection of the correct entry point critically affects the fetoscope maneuverability and the possibility to burn all the anastomoses.

quantitative ultrasound fetal lung maturity analysis (quantusFLM^{TM}) was created through 2 different modules. The first feature extractor module combined invariant features to geometric and photometric transformations. Then, lung classification algorithms such as regression models, trees and neural networks were trained on the second module to identify high and low risks of respiratory morbidity. The performance achieved an ACC = 87%, SEN = 91%, and SPE = 86%. To the best of our knowledge, this is the first work reporting blind validation of quantitative imaging analysis software designed specifically to predict neonatal respiratory morbidity. Its main difference between previous published studies is the large number of images utilized. Nevertheless, three main limitations arise: 1) a single cohort was used, 2) the image acquisition and delineation were performed by highly-trained personnel, and 3) the sample size prevented the performance evaluation within narrow gestational age ranges.

4. Examples of fetal surgery applications

There have been attempts to implement segmentation and classification algorithms for the aforementioned organs in real surgeries at pre-operative stage. In the following, examples of fetal diseases for which these scientific approaches have been applied (or could be employed in the near future) at the pre-operative level are highlighted.

4.1. Twin-to-twin transfusion syndrome

Twin-to-twin transfusion syndrome (TTTS) is a rare, potentially fatal condition that affects around 10–15% of monochorionic twins. It originates from the blood vessels (small anastomoses) located inside and on the surface of the shared placenta connecting both babies (Valsky et al. (2011)). In this syndrome, blood is not evenly distributed, but part of it is rather diverted from one of the twins (the *donor*) to the other one (the *recipient*). Fetoscopic laser photocoagulation therapy is one of the most effective TTTS treatments (see Fig. 10) and it consists in closing blood vessels connecting the twins (Slaghekke et al. (2014)). If all blood vessels linking the babies are efficiently burned, the procedure only needs to be performed once. Otherwise, to reentry in a different point is highly associated with fetuses' mortality (Slaghekke et al. (2014)). Therefore, the selection of the correct entry point is probably the most critical factor of the intervention as it directly affects the fetoscope maneuverability and the possibility to reach all the anastomoses.

Although all methods described in Section 3.3.1 and 3.3.2 can be used to effectively know the placenta position, only Ceresa et al. (2017a) proposed a surgical software for planning the insertion trajectory taking into account the specific placenta and blood vessels localization. In their simulated surgical scene, the authors load a model of a 2.0 mm fiber fetoscope and let the surgeon inspect the insertion point interactively using a replica of the fetoscope connected to an NDI Aurora tracker. Previously, Luks et al. (2001) studied the anatomy of each amniotic cavity, the point of placental insertion of the umbilical cord and the location of the inter-twin membrane in relation to the port placement. Three models were used to calculate the optimum port entry point, as well as the length and angle of the curved tool required to reach the target region at the desired angle. In both works, the 3D reconstruction rendered enhanced anatomical understanding at the preoperative stage, and allowed manipulation of the imaging to facilitate surgical planning. Moreover, Liao et al. (2009) developed a fast image mapping system to fuse endoscopic image mosaics with 3D US for assisting directly the intrauterine laser photocoagulation therapy. A Polaris Vicra optical tracking system was employed to track the endoscope position and the US probe. The endoscopic images were registered to the placenta (phantom) surface extracted from the 3D US through a GPU-based multi-image processing approach. The image mapping accuracy was 2.8 mm on average.

4.2. Congenital heart disease

As explained in Section 3.4, CHD refers to a structural heart abnormality present at birth. The most common surgical treatment of CHD in the fetus is performed as follows. A 18- to 19-gauge needle is introduced into the fetal heart and the ventricle pointing toward the corresponding valve through the uterus under continuous US guidance. After ventricular puncture, the needle is positioned with the tip within the outflow tract below the valve. The trocar is removed and the wire with the preloaded and measured balloon catheter is inserted into the needle and advanced through the valve. The balloon is positioned across the valve on the basis of previous measurements and appearance on US. When the position is confirmed, one or more inflations are accomplished.

Although no sophisticated systems have been implemented to plan the pre-operative stage of this surgery, some preliminary works recently (Sridevia and Nirmala (2016); Nirmala and Sridevi (2016); Bridge et al. (2017)) address this new line of research. Therefore, the aforementioned methodologies could be exploited to design a useful pre-operative planning framework.

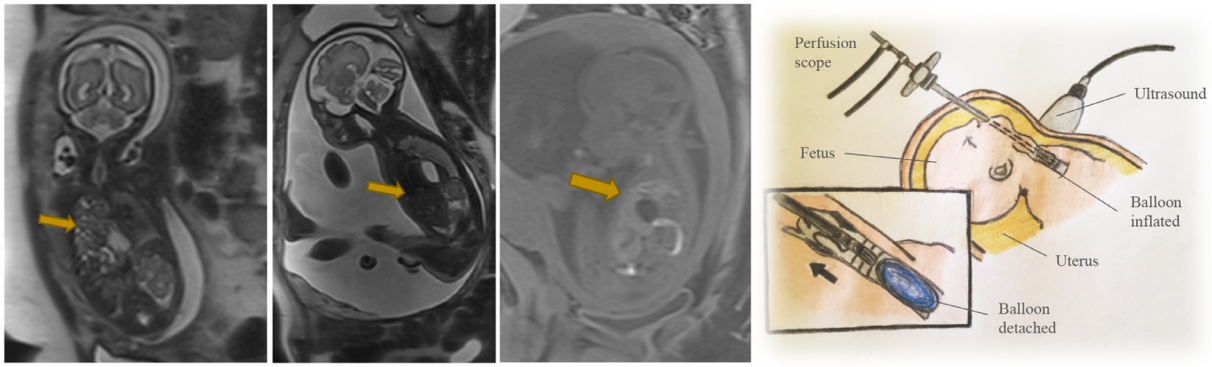

Fig. 11. Congenital diaphragmatic hernia (*from left to right*): Coronal TruFISP, sagittal HASTE and T_1-weighted images showing a left diaphragmatic defect through which the stomach and loops of bowel are herniated (*first figure*). A partly herniated T_2 hypointense liver located anterior in the thorax is also presented (*second figure*). Furthermore, a herniated T_1 bright liver is shown (*third figure*). The fetal endoscopic tracheal occlusion is finally illustrated (*fourth figure*).

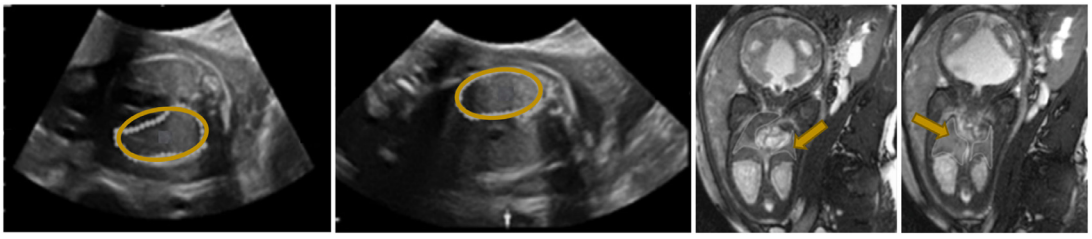

Fig. 12. Pulmonary hypoplasia (*from left to right*): US with bilateral large echogenic kidneys, US with a large posterior fossa cyst, Coronal single shot MRI T_2 images of fetal brain, chest and upper abdomen that show large posterior fossa cyst communicating with the 4th ventricle, large hyperintense kidneys. The lungs depict a lower than normal volume close to that of surrounding muscles.

4.3. Congenital diaphragmatic hernia

Congenital diaphragmatic hernia (CDH) is a birth defect occurring because the diaphragm is not completely formed (see Fig. 11). Since there is a hole in this structure, the organs in the abdomen (stomach, bowels, liver) may ascend to the chest, thereby compressing the lungs and preventing their normal development (pulmonary hypoplasia). The intervention consists in a minimally invasive therapy called Fetal Endoscopic Tracheal Occlusion (Mesas-Burgos et al. (2017)). Through a small incision in the skin (about 3 mm) a tiny fetoscope is inserted and under direct vision the endoscope is advanced through the mouth of the fetus down to the trachea, where a balloon is inflated. This balloon is left inflated for a period ranging from 3 to 6 weeks and then it is removed. The rest of the pregnancy can be managed in a usual way (Terui et al. (2017)).

To the best of our knowledge, no works have been reported on the design of a pre-operative planning system to aid CDH fetal surgery. Nevertheless, the methodology presented by Keraudren et al. (2015) could be employed to compute the abnormal displacement of the fetal lungs and liver with respect to their usual localization. Once surgeons know the position deviation of these organs, they could estimate approximately the volume of the inflated ballon to prevent the compression of the lungs.

4.4. Pulmonary hypoplasia

Pulmonary hypoplasia (PH) is a condition in which the lungs are abnormally small, and do not have enough tissue and blood flow to allow the fetus to breathe on its own (see Fig. 12). Most often the PH is a secondary problem due to another diagnosis such as CDH or bladder outlet obstruction. Either something is compressing the lungs or preventing the lungs from developing properly (Matturri et al. (2003)). Its diagnosis is based upon measurements (volume and size) taken from US and MRI images in order to determine whether the problem is life-threatening for the baby (Wu et al. (2017)). Fetal surgery can help reverse the effect of the pH while the baby is still developing in the womb.

No research has been carried out in the field of pH surgical planning. However, if pH appears after a CDH abnormality its pre-operative planning could be directly related. A second workflow in the CDH planning framework could be added by taking advantage of the algorithms reviewed in Section 3.5. In this line, a wider analysis of relevant pH risk factors such as size, position and maturity of fetal lungs could be reported on both US and MRI image modalities to deal with successful surgeries.

4.5. Pulmonary sequestration

A pulmonary sequestration (PS) is a cystic piece of abnormal lung tissue that appears next to or within the lung. The diagnosis is made by prenatal US findings of an echogenic (bright) mass appearing in the fetus chest (see Fig. 13). The mass may displace the heart from its normal position or push the diaphragm downward, but the key feature of a sequestration is the finding of an artery leading from the cystic mass directly to the aorta (Yoon et al. (2017)). If the fetus has a very large lesion that will make resuscitation after delivery dangerous, a specialized delivery can be planned called the EXIT (ex utero intrapartum treatment) procedure.

Once again, there is no literature on the topic of PS surgical planning. Although several follow-up methods proposed in Sections 3.4 and 3.5 could be utilized to detect the echogenic mass through the localization of the fetal heart and lungs, its full implementation has not been addressed before. The automatic segmentation of the large vessel arising from the descending

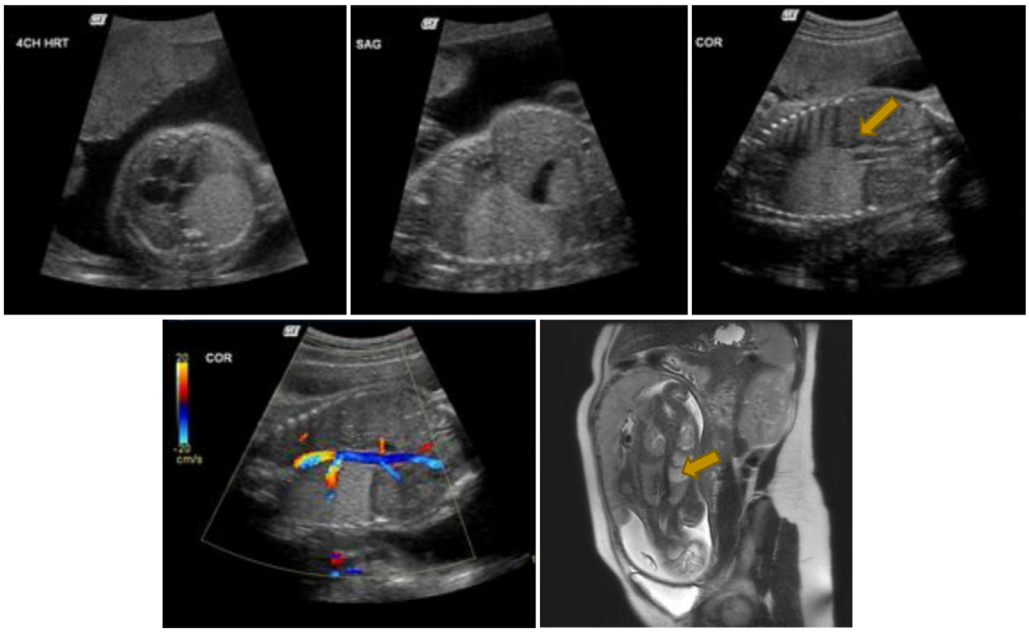

Fig. 13. Pulmonary sequestration (*from left to right*): US axial, US sagittal, US coronal, US coronal Doppler and MRI T_2 SSFSE. *US images*: Large hyperechoic homogeneous mass is identified in the left chest pushing the heart completely into the right chest. In addition, a large feeder vessel is identified arising from the descending thoracic aorta. *MRI image:* T_2 hyperintense mass in left chest with flow void arising from the descending thoracic aorta.

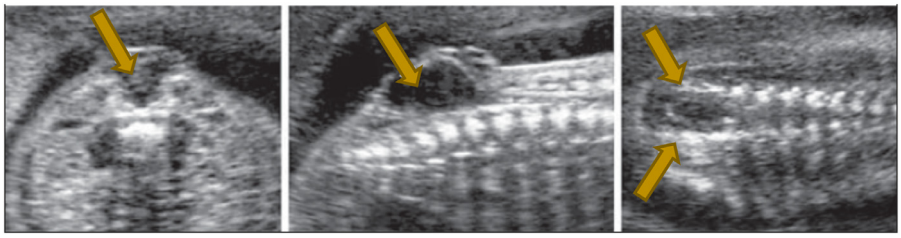

Fig. 14. Figure extracted from Fichter et al. (2008): *Axial* (*left*), sagittal (*middle*) and coronal (*right*) view planes of US scans show a SB defect. The cystic lesion is characterized by a fluid-filled sac, where the neural placode is lifted out of the spinal canal and hair-like nerves can be seen rising through the sac (sagittal view).

thoracic aorta remains unexplored, despite that it plays a crucial role in the pre-operative planning. In this case, 3D US doppler serves to efficiently detect this artery.

4.6. Spina bifida

Spina bifida (SB) or *Myelomeningocele* is a birth dysraphic defect where an incomplete closing of the backbone and membranes exists around the spinal cord, leading to severe impairment throughout the patient's lifetime (Fichter et al. (2008)). SB is caused by a neurulation failure in the 4^{th} week of gestation. It can be diagnosed prenatally in over 80% of all cases by maternal serum α-fetoprotein and obstetrical US (see Fig. 14). Indirect signs such as irregularities of the bony spine or a bulging within the posterior contour of the fetal back are present in most SB cases before the 24^{th} week of gestation.

The SB fetal surgery implies a delicate procedure where surgeons open the mother uterus and close the opening in the baby's back while they are still in the womb. The first cases of SB repair were performed fetoscopically to minimize maternal risks, but proved unsuccessful due to small-calibered instruments, anterior placenta, unfavorable fetallie, maternal obesity, technical difficulties of the surgical procedure, among others (Izatt et al. (2007)). Recent advances in robotics have solved some of these technical problems, but an accurate surgical planning is still required to increase the success rate. Although no works on SB surgical planning have been reported so far, Kohl et al. (2012) performed a pre-operative sonoanatomic examination of SB aperta to allow prediction of surgical complexity during subsequent minimally-invasive fetoscopic closure on 9 human fetuses. In this way, authors made a pre-operative assessment to collect essential information for facilitating surgical planning of the procedure.

5. General performance assessment

Tables 2–5 summarize the validation done on the 123 articles discussed in Section 3. They involve all fetal image segmentation and classification methods in terms of the fetal organs here presented, namely brain, placenta, lungs and liver (thorax), heart, and the whole fetus.

KEY TO TABLE

Modality: MRI (Magnetic Resonance Imaging), US (Ultrasound), and 3D US (3-dimensional Ultrasound).

Automation (Auto): S (Segmentation), C (Classification), + (Full), - (Interactive guidance), and = (Substantial guidance).

Evaluation (Eval): QL (Qualitative), and QN (Quantitative with manual delineations / clinical reference or standard).

Validation (Val): Number of subjects (number of images).

From Tables 2–5, some general comments on the state-of-art reviewed can be made.

- In the field of medical imaging, some simulation, phantom, and clinical studies are usually adopted to validate segmentation and classification approaches of anatomical organs. Such clinical studies are particularly based on the use of manual delineation as the *gold standard* to be correlated with clinical outcome. Simulation and phantom studies have been carried out, but essentially clinical data is utilized during the validation process. This is due to the fact that genuine simulations and phantoms are hard to define.
- Manual delineation on clinical data is the preferred modus of performance assessment. However, the inter-expert and intra-expert variability in the delineation procedure makes it more sensible to outliers. It is therefore important to be careful during the results interpretation when manual tracing is utilized as the main reference clinical outcome.
- There are no guidelines in the performance evaluation. Although authors use well-known and common measures, a direct comparison of methods cannot be easily performed as a whole, since each work is validated through different measures.
- From Tables 2–5 we can see that 55 articles were validated on databases with less than 60 images (44.7%), and 11 articles did not specify the number of images used. Hence, a higher degree of robustness is needed in forthcoming works.
- There exists a lack of standardization in the data collected. There are no public databases on which different research groups can compare methods. Thus technical benchmarks can only be performed between algorithms developed by individual institutions or through research collaborations. This problem is aggravated in US as there is a huge variability in its low-quality in comparison to other image modalities such as MRI.

6. Discussion and future prospects

Fetal MRI has gained wide acceptance in clinical practice as a valuable tool to complement prior prenatal (3D) US findings. Automatic segmentation and classification algorithms can nowadays facilitate the diagnosis of *in utero* abnormalities of fetal organs. In addition, thanks to the major breakthroughs in medical equipments and the recent advancements in the image acquisition protocols, more scalable and high-quality population reference studies can be performed. Crucial information is obtained to aid the decision-making process of a specific treatment. Therefore, the accuracy and quality of preoperative fetal MRI and US are of the utmost importance to avoid an inaccurate diagnosis, since it could lead to suboptimal or inappropriate intervention. Fetal imaging needs to be differently approached in comparison to adult organs literature due to the complex and rapid changes that all fetuses undergo during the gestational period.

Fig. 15 describes the intrinsic relation between the reviewed techniques, organs and image modalities. It gives an overview of which are the most employed algorithms, which of them have been independently applied to different organs and image modalities, and which research lines regarding fetal imaging remain unexplored.

Although new MRI classification approaches are capable to segment fetal organs robustly, the final segmentation is still subject to inter-slice motion artifacts. High-level features (e.g., Fig. 15 - first row) are widely used to overcome the low contrast and weak boundaries, which help to obtain a better description of all organs and background than intensity-based low-level features. Likewise, textural features are also utilized in (3D) US data to deal with fetal organ analysis such as placenta classification or lungs segmentation. However, these features have a longer runtime. The trend is to improve the robustness by adding other features such as geometry and morphology. Moreover, deep learning tech-

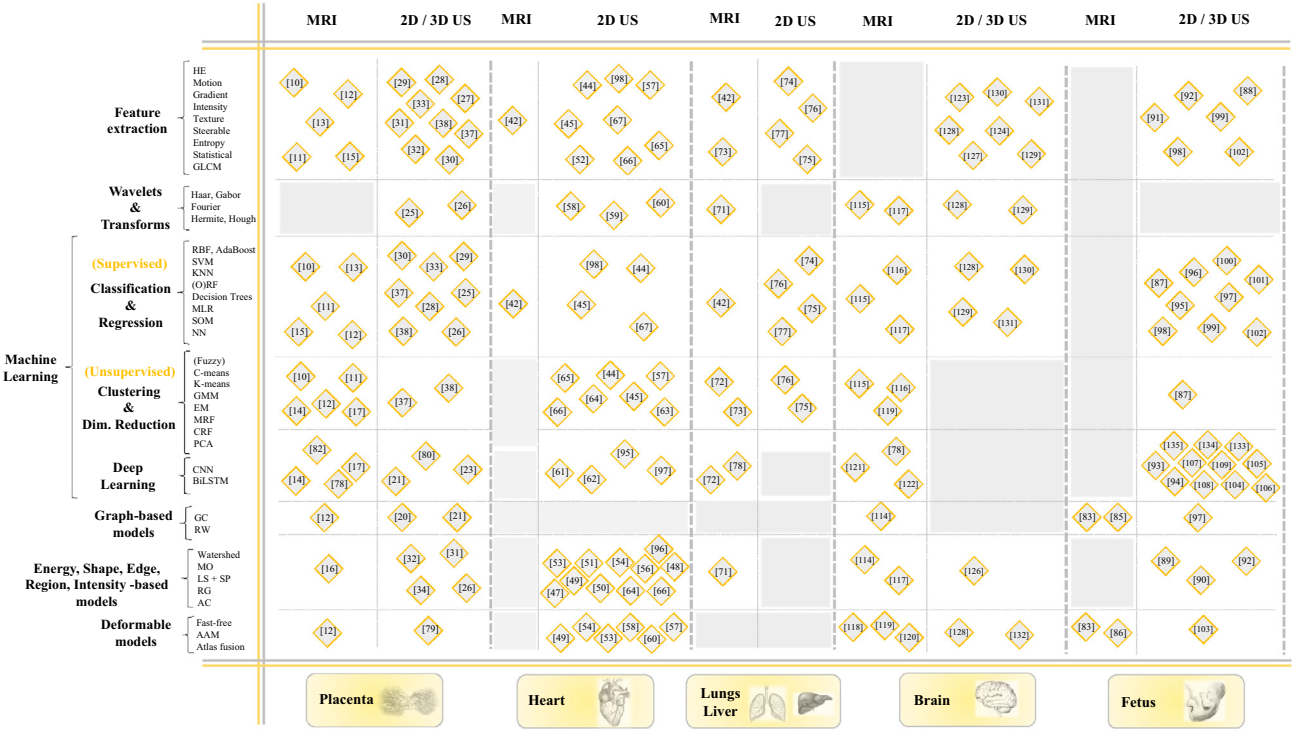

Fig. 15. Categorization of the reviewed techniques according to each fetal organ and image modality. The nomenclature and the references in square brackets used are the same as defined in Tables 2–5.

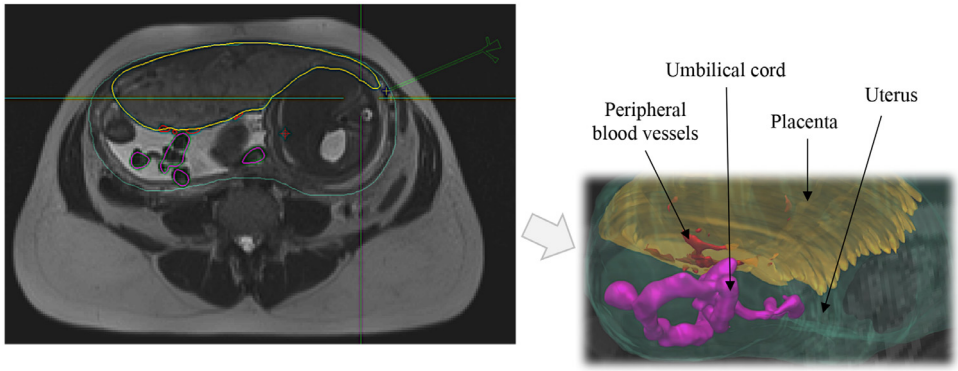

Fig. 16. A 3D model of the mother uterus: Axial MRI view (*left*), and 3D isosurface (*right*): placenta (*yellow*), peripheral blood vessels (*red*), uterus (*green*) and umbilical cord (*purple*). A green simulated fetoscope is located on the *top-right* of the Axial view. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

niques (e.g., Fig. 15 - fifth row) can extract robust features using a network that takes into account local and larger contextual features at the same time (multi-scale features). Despite its recent advances, deep learning has not been applied yet to fetal heart and lungs segmentation / classification in MRI and 2D US, respectively. Recently, these approaches have also incorporated CRF to address the disconnectivity of labels resulting from prediction. The application of a CRF as a recurrent neural network might be a solution for back-to-back training parameters involved in each algorithm iteration. Nevertheless, the impact of high-level features and CRF in fetal imaging has not been investigated in detail (e.g., Fig. 15 - fourth row). Other techniques such as random forests and support vector machines (e.g., Fig. 15 - third row), or active contours and morphological operations (e.g., Fig. 15 - seventh row) are widely employed in placenta, lungs, liver and heart. A complete 3D segmentation could be used to avoid the accumulated propagation error in all organs, which is typically higher in terminal slices due to a large shape change of the organs between slices. Researchers should also take on board super-resolution methods to reconstruct all MRI views to improve the final results and reduce the noise.

In the case of the placenta, several feature extractors, supervised and unsupervised techniques have been successfully applied in both MRI and 2D / 3D US (e.g., Fig. 15 - first and second columns). Therefore, researchers can go a step further and improve the anatomical understanding of the placenta by rendering a 3D model of its vasculature through the fusion of MRI and US Doppler. Likewise, MRI and 3D US could be simultaneously used to locate the vessels. In MRI, vessels appear as black dots whilst in 3D US they are seen as black lines and cavities adhered to the placenta boundary. For instance, vessels segmentation in TTTS has to be carried out between approximately 16–26 weeks of gestation to prevent the death of both fetuses. The exact location of the umbilical cord insertions should be also addressed in the near future as the vasculature emanates from it. All the aforementioned improvements would allow surgeons to perform better in the preoperative stage and carry out self-reliant laser treatment operations.

Non-invasive diffusion-weighted MRI (DW-MRI) is recently emerging to quantify *in vivo* global placental perfusion without the need for exogenous contrast agents, which has a crucial role for the effective delivery of nutrients and gases to the fetus. More specifically, the intra-voxel incoherent motion (IVIM) modality estimates the fraction of moving blood in the total placenta volume under the assumption that the DW signal is associated with both water diffusion in tissues and blood diffusion inside capillaries. For instance, Moore et al. (2000) toke advantage of these properties to empirically segment the placenta into fetal (chorionic plate) and maternal (decidual plate) compartments. The intervillous placenta space was hard to divide due to the complex and hierarchical villous branches. This compartment differentiation was also tackled in Kording et al. (2015) by applying a K-means clustering of the time-to-peak values obtained from dynamic contrast-enhanced (DCE) MRI. An adapted gamma function was fitted to the discrete concentration time curves to improve the perfusion and segmentation results. A mean DC of 0.91 $\pm$ 0.08 and 0.89 $\pm$ 0.08 was found for both compartments. Freshly, You et al. (2017) used *in vivo* IVIM to semi-automatically detect the compartments via GrowCut multi-labeling. The perfusion fraction was 0.58 $\pm$ 0.11 and 0.56 $\pm$ 0.13 in the fetal and maternal compartments, respectively. Great efforts are being made to enhance the DW-MRI. Flow and fluid measurement methods are confounded by the interaction of relaxation and perfusion features, which are rarely simultaneously considered. In this framework, Melbourne et al. (2016) proposed a three compartment model comprising fast and slowly circulating fluid pools motivated by the different flow characteristics across the placenta. The placenta pool was fitted to overlapping multi-echo T2 relaxometry and IVIM.

Regarding placenta classification in 2D US, it has been shown that calcification is higher in placentas with *Grannum* grades of II and III. Early preterm calcification is associated with a higher incidence of adverse maternal and fetal outcomes. Grade III is often the most confusing stage in placenta maturity analysis due to its subjective visual characteristics. Encoding features as a function of the *Grannum* grade could improve grade III detection (e.g., Fig. 15 - first row and second column). Dense sampling could help as well, since it generates a high number of samples to discriminate between different grades. Nevertheless, the *Grannum*'s method lacks objectivity, precision and reproducibility. Indeed, it is presently unused in the clinics. Therefore, suitable alternatives are needed. In particular, an objective and reproducible quantitative grading system for the assessment of 2D US images should be designed. Vector methodologies such as vector of locally aggregated descriptors and Fisher vector could be implemented as they demonstrated better classification results. A multi-class SVM could improve the classification of these features.

Energy, shape, edge, region and intensity -based models are already implemented for the heart in 2D US (e.g., Fig. 15 - seventh row and fourth column), but scarcely used in lung, placenta and liver. In particular, level set approaches are often used in heart segmentation. The shape prior constrains the propagation to a predefined shape in areas of ambiguity such as missing chamber walls or spurious edges yielding a better segmentation agreement. Re-

searchers have easily adapted shape prior to different snake algorithms when the shape is unknown. In addition, deformable models (shape + texture) are widely used to detect the fetal heart structures (e.g., Fig. 15 - eighth row and fourth column). The trend is to utilize the heart structure and the motion information simultaneously. This can be achieved by using US video sequences instead of single images. Although other organs do not take advantage of fuzzy clustering techniques, in fetal heart they are becoming increasingly important (e.g., Fig. 15 - fourth row and column). Pre-operative surgical systems to aid fetus cardiomegaly, congenital heart disease or defect, prenatal ventricular septal defect have been scarcely reported and therefore there is room for improvement. Likewise, 3D/4D annulus fetal heart segmentation deserves to be researched in the fetal domain. In adult hearts, segmenting the mitral valve during closure and throughout a cardiac cycle from 3D/4D US has been demonstrated essential for creation and validation of mechanical models, improved visualization and understanding of mitral valve behavior, and pathology diagnosis. Often, mitral valve segmentation from 3D/4D US images either requires extensive user interaction and initialization, does not maintain the valve geometry across a cardiac cycle, or cannot produce a detailed coaptation line and surface (Schneider et al. (2011, 2012)).

Segmentation of fetal lungs on US images should be addressed. Only one work has been published so far. Certainly, it is an extremely challenging task due to the low signal-to-noise ratio. Currently, no automatic method exists. In the clinics, fetal lung volumetry is assessed by manually segmenting individual US slices at a postmortem stage or individual MRI motion corrupted slices. On the other hand, it would be appealing to improve the state-of-the-art techniques that provide the textural description of fetal lungs as a function of the gestational age (e.g., Fig. 15 - first + third rows and sixth column). This would enable the introduction of non-invasive techniques into clinical practice to test fetal lung maturity. For example, neonatal respiratory morbidity could be predicted based on quantitative analysis, thus replacing current tests involving amniotic fluid.

The clinical applications of obstetric ultrasonography can be extended with an improved understanding of the fetal liver anatomy, which is mostly overlooked in the US literature, in particular its morphology. Hepatic calcifications in the fetus are hyperechogenic areas that are detected by US imaging. They are termed fetal liver calcifications. The incidence of fetal hepatic calcifications is undetermined. Although isolated liver calcifications are relatively common, their clinical significance and management are not precisely known. Calcifications can form either at the liver parenchyma or at the liver surface. Fetal liver calcifications were found to be associated with other abnormalities on ultrasonographic examination, such as ventriculomegaly, ventricular septal defect, ascites, and echogenic gut. Possible causes for these lesions have been associated with infection, ischemic insults, portal and hepatic vein thromboemboli, tumors, chromosomal abnormalities, and sludge or lithiasis in the gallbladder.

This review demonstrates that new technologies and segmentation / classification algorithms are currently being investigated to aid the fetal preoperative stage, particularly in specialities where surgery is complicated, and where the pathology is unusual. Great efforts have been made to produce high-fidelity simulators for fetal interventions in spite of being a relatively young research field, but there is still a long way to go. The research in other fields like brain, breast cancer or pulmonary nodule segmentation is more mature, partly because machine learning and deep learning methods were implemented earlier in time.

7. Conclusion

This work provides the most recent landmarks on segmentation and classification methodologies in the context of fetal imaging. A total of 123 relevant studies in which state-of-the-art methods are used to analyze fetal structural anatomies to identify possible abnormalities or complications have been surveyed. Although fetal MRI and (3D) US are still improving, with higher resolution and enhanced signal-to-noise ratio, new processing approaches are emerging. Given the constraints that intrauterine environments impose and the intricate fetal anatomy, not only well-designed computer-assisted frameworks are required, but also novel, fast, and efficient ad hoc fetal algorithms are needed. Fetal surgery interventions, for instance, are currently taking benefit from such techniques to achieve an enhanced pre-operative planning. This review grasps the most recent computer-based contributions to the field of fetal lungs, liver, heart and placenta imaging.

Acknowledgements

This work was supported by CELLEX foundation and the Google Women Techmakers scholarship awarded to Jordina Torrents-Barrena. Also this work was funded by the Spanish Ministry of Economy and Competitiveness under the María de Maeztu Units of Excellence Programme [MDM-2015-0502].

Table 2
Part I - Summary of all works (123 articles) included in this review.

Article	Modality	Organ	Method	Auto	Performance		Val
					Eval	*Value*	
Anquez et al. (2010) [83]	MRI		Skeleton + Articulated model + Graph cuts	S(-)	QN	Kappa = 0.89	24(24)
Bibin et al. (2009) [84]	MRI 3D US		Segmentation + Contour regularization + Mesh surface reconstruction + Model integration	S(+)	QL	-	5/4(5/4)
Zhang et al. (2016) [85]	MRI		Graph-based	S(+)	QN	DC = 0.69 ± 0.08	10(10)
Zito et al. (2013) [86]	MRI		Free-form ROIs + Barke equation	S(+)	QN	Error = 1.96%	20(20)
Gupta et al. (2011) [87]			FCM + SVM + CRF	S(-)	QL	-	2(2)
Krishnan et al. (2014) [88]			Anisotropic diffusion + Iso-intensity / edge-focusing contouring	S(+)	QN	Error = 4.8%	-
Anquez et al. (2013) [89]		Whole	Bayesian formulation + Level-set + Rayleigh distribution	S(-)	QN	DC = 0.89	19(19)
Choong et al. (2012) [90]			Variational level-set	S(-)	QN	Error = 2.38%	-
Dahdouh et al. (2013) [91]			Legendre moments + Shape envelope	S(-)	QN	DC = 0.72	16(16)
Dahdouh et al. (2015) [92]			Shape-constrained level-set + Statistical intensity priors + Legendre moments	S(-)	QN	DC = 0.7	18(18)
Li et al. (2017) [93]			CNN + Transfer learning	S(+)	QN	ACC = 79.3% *	4(601)
Yang et al. (2017b) [94]			3D CNN + BiLSTM	S(+)	QN	DC = 0.8 *	104(104)
Carneiro et al. (2008a) [95]			Probabilistic model + Discriminative / generative classifiers	C(+)	QL	-	40(40)
Rahmatullah et al. (2011) [96]			AdaBoost	C(+)	QN	ACC = 76.29%	30(30)
Zhang et al. (2012) [97]	US	(gestational sac)	Two AdaBoost + Multi-scale cut	C(+)	QN	Error = 7.5 ± 5%	31(31)
Yaqub et al. (2014) [98]		Whole	Regression RF + Geometrical features	C(+)	QN	Graphical	87(87)
Yaqub et al. (2015) [99]		Whole	RF + Translation / orientation invariant features	C(+)	QN	ACC = 83% *	30K(30K)
Ni et al. (2014) [100]		(abdominal)	Radial components + RF + Selective search	C(+)	QN	ACC = 85.6%	223(223)
Ni et al. (2013) [101]		(abdominal)	Radial components + AdaBoost + Selective search	C(+)	QN	Graphical	100(100)
Lei et al. (2015a) [102]		(facial)	SVM + FV + Dense root scale invariant / spatial transform features	C(+)	QN	ACC = 93.27%	(1735)
Liu et al. (2012) [103]		(head)	AAM + Template matching	C(-)	QN	Error = 28%	51(51)
Chen et al. (2015b) [104]		(abdominal)	CNN + Transfer learning	C(+)	QN	ACC = 89.6%	219(8718)
Chen et al. (2014) [105]		(abdominal)	CNN + Specific ROIs + Transfer learning	C(+)	QN	ACC = 90.7% *	219(8718)
Yu et al. (2018) [106]		(facial)	CNN + Transfer learning	C(+)	QN	ACC = 96.98%	(4849)
Baumgartner et al. (2016) [107]		(heart)	CNN + Saliency maps	C(+)	QN	ACC = 69%	1003(1003)
Chen et al. (2015a) [108]		(abdominal, facial, heart)	CNN + LSTM + Transfer learning	C(+)	QN	ACC = 88%	631(631)
Baumgartner et al. (2017) [109]		Whole	Two-fold CNN	C(+)	QN	DC = 0.798	2694(2694)
Schlemper et al. (2018) [133]		Whole	Baumgartner et al. (2017) + Attention mechanisms	C(+)	QN	ACC = 98%	2694(2694)
Li et al. (2018a) [134]		(brain)	CNN + Two loss functions	C(+)	QN	Error = 3.45 *mm*	72(72)
Li et al. (2018b) [135]		(brain)	Iterative transformation network	C(+)	QN	Error = 3.82 *mm* *	72(72)

Table 3
Part II - Summary of all works (123 articles) included in this review.

Article	Modality	Organ	Method	Auto	Performance		Val
					Eval	*Value*	
Studholme et al. (2011) [110]			*Review*	S(+/-)	QN / QL	-	-
Studholme et al. (2014) [111]			*Review*	S(+/-)	QN / QL	-	-
Benkarim et al. (2017) [112]			*Review*	S(+/-)	QN / QL	-	-
Makropoulos et al. (2018) [113]			*Review*	S(+/-)	QN / QL	-	-
Anquez et al. (2009) [114]			Template matching + Morphological operations + Biometrics + Graph cuts	S(-)	QN	Kappa = 0.89	24(24)
Ison et al. (2012) [115]			RF + 3D Haar features + MRF	S(+)	QN	ACC = 72.7% *	43(43)
Keraudren et al. (2014) [116]	MRI	Brain	SIFT + K-means + SVM + RF + CRF	S(+)	QN	DC = 0.93	66(66)
Kainz et al. (2014a) [117]			RF + 3D Gabor features + 2D level-set	S(+)	QN	DC = 0.90	50(50)
Taimouri et al. (2015) [118]			4D atlas fusion + Template matching	S(+)	QN	ACC = 94%	30(366)
Wright et al. (2014) [119]			Spatiotemporal atlas + EM + MRF	S(+)	QN	Error = 1 mm	80(80)
Tourbier et al. (2015) [120]			Multiple atlas + Normalized correlation + Global weighting	S(+)	QN	DC = 0.90	12(12)
Khalili et al. (2017) [121]			2D multi-scale CNN	S(+)	QN	DC = 0.93 *	30(30)
Salehi et al. (2018) [122]			Auto-context fully CNN	S(+)	QN	DC = 0.88 *	35(35)
Vélasquez-Rodríguez et al. (2015a) [123]	US	(cerebellum)	Matching cubes + Spherical harmonics + Gray-level objective function	S(-)	QN	DC = 0.75	10(10)
Vélasquez-Rodríguez et al. (2015b) [124]	US	(cerebellum)	Matching cubes + Spherical harmonics	S(-)	QN	DC = 0.69	10(10)
Becker et al. (2010) [125]	US	(cerebellum)	3D point distribution model + Genetic algorithm	S(+)	QL	-	5(5)
Gutiérrez-Becker et al. (2013) [126]	US	(cerebellum)	3D statistical shape model + Nelder-Mead	S(-)	QN	DC = 0.8	20(20)
Benavides-Serralde et al. (2009) [127]	US	(4 structures)	VOCAL	S(-)	QN	Intra-class correlation = 0.72 *	78(78)
Yaqub et al. (2013) [128]	US	(4 structures)	RF + Template deformation + Hough transform + Statistical features	S(-)	QN	ACC = 74.5% ± 11.25 *	51(51)
Yaqub et al. (2012) [129]	US	(4 structures)	Discriminative RF + Haar, cuboid, binary, unary features	S(+)	QN	ACC = 91.95% *	40(40)
Cerrolaza et al. (2017) [130]	3D US	(skull)	RF + Semantic, structural features + Geodesic distance	S(+)	QN	DC = 0.8 ± 0.03	59(59)
Carneiro et al. (2008b) [131]	US	(head)	AdaBoost + Statistical features	S(+)	QL	-	1426(1426)
Cuingnet et al. (2013) [132]	3D US	(skull)	Ellipsoidal shell model + Deformable template	S(-)	QN	5.1 ≤ Error ≥ 5.8 mm	78(78)
Kainz et al. (2014) [70]	MRI	Thorax	TV minimization + Fast rotation invariant spherical harmonics + RF + Logistic regression + Geodesic AC + 3D multi-scale vessel enhancement + Opening	S(-)	QN	DC = 0.87, FPR < 10%	29(29)

Table 4
Part III - Summary of all works (123 articles) included in this review.

Article	Modality	Organ	Method	Auto	Performance		Val
					Eval	*Value*	
Wang et al. (2015a, 2015b) [10],[11]			RF + CRF + High-level features	S(-)	QN	DC = 0.79 ± 0.07	6(6)
Wang et al. (2016b) [12]			Wang et al. (2015a, 2015b) + 4D Graph Cut + Fast Free-Form Deformation	S(-)	QN	DC = 0.82 ± 0.02	16(16)
Wang et al. (2016a) [13]			ORF + DyBa + Tree Growing + Shrinking + Gray-level features	S(-)	QN	DC = 0.89 ± 0.04	16(16)
Alansary et al. (2016) [14]	MRI		CNN + Dense CRF + PVR + mean-curvature flow skeleton	S(+)	QN	DC = 0.72 ± 0.11	66(66)
Ceresa et al. (2017a,b) [15]			Super-resolution + local features + SVM	S(+)	QL	-	-
Miao et al. (2017) [16]			Alansary et al. (2016) + Morphological operations	S(+)	QL	Survey	4(4)
Wang et al. (2018) [17]			Interactive CNNs + Geodesic distance maps + CRF	S(-)	QN	DC = 0.89 ± 0.05	25(25)
Torrents-Barrena et al. (2018) [82]		(umbilical cord)	3D CNN + LSTM	S(+)	QN	DC = 0.79 ± 0.03	10(10)
Stevenson et al. (2010) [19]	3D US		-	S(-)	QN	Statistical	10(10)
Stevenson et al. (2015) [20]	3D US		Weighted Random Walker	S(-)	QN	DC = 0.87 ± 0.1 *	58(58)
Oguz et al. (2016) [79]	3D US		Multi-atlas label fusion	S(+)	QN	DC = 0.83 ± 0.05	13(13)
Looney et al. (2017) [21]	3D US		Random Walker + *DeepMedic* by *Kamnitsas et al. (2017)*	S(+)	QN	DC = 0.73 (0.66, 0.76)	300(300)
Yang et al. (2017) [23]	3D US		3D fully CNN + Hierarchical BiLSTM	S(+)	QN	DC = 0.643	104(104)
Looney et al. (2018) [80]	3D US		Fully CNN	S(+)	QN	DC = 0.81	2393(2393)
Linares et al. (2003) [28]		Placenta	Texture features + C5.0	C(+)	QN	ACC = 87%	12(12)
Linares et al. (2004) [29]			Texture features + relief-F + KNN	C(+)	QN	ACC = 60.71%	101(101)
Ryan et al. (2008) [34]			Flood-filling + Thresholding + *Bresenham*'s line + RG	C(-)	-	-	100(100)
Malathi et al. (2009) [25]			Stereo mapping + (Haar) wavelet + NN	C(+)	-	-	250(250)
Liu et al. (2009) [33]			Gray-level features + multi-class SVM + Super-ball / Ultra-radius	C(+)	QN	Recognition rate = 92%	200(200)
Malathi and Shanthi (2010a, 2010b) [31],[32]	2D US		Malathi and Shanthi (2010c) + RG	C(-)	QL	Visual	5(5)
Malathi and Shanthi (2010c) [27]			HE + Texture features	C(-)	QN	Statistical	-
Malathi and Shanthi (2011) [26]			Stereo mapping + (Haar) wavelet + Watershed + NN	C(+)	QL	Visual	5(5)
Chen et al. (2011) [30]			Texture features + MLR	C(-)	QN	MLR correlation	151(151)
Moran et al. (2013) [35]			Thresholding + Flood-filing + Quantification	C(-)	QN	AUC = 0.628	90(90)
Li et al. (2014), Lei et al. (2015) [37],[39]			Dense sampling + DAISY + VLAD + GMM + PCA + K-means + SVM	C(+)	QN	ACC = 0.874, SEN = 0.996, SPE = 0.874	311(311)
Lei et al. (2014, 2015) [38],[39]			LIOP features + Harris-Laplace + FV + GMM + K-means + PCA + SVM	C(+)	QN	ACC = 0.9375, SEN = 0.9804, SPE = 0.9375	443(443)
Chikop et al. (2014) [71]	MRI	Lungs	AC + Morphological operations + Haar wavelet + Fourier transform	S(-)	QL	Graphical	6(6)
Keraudren et al. (2015) [42]	MRI	Lungs, Liver, Heart	RF + Steerable features	S(+)	QN	ACC > 90%	55(55)
Rajchl et al. (2017) [72]	MRI		CNN + dense CRF	S(+)	QN	ACC = 74.9 ± 6.7%	55(55)
Wang et al. (2018a) [78]	MRI	Lungs, Brain, Placenta	Coupled CNN	S(-)	QN	0.85 ± 0.06 ≤ DC ≥ 0.96 ± 0.02	18(18)
Li et al. (2010) [73]			GLCM features + GMM + EM + Morphological thinning + Vector field convolution	S(-)	QL	Graphical	-
Prakash et al. (2002) [74]	2D US	Lungs	Texture features + (NN, KNN, RBF, SVM)	C(+)	QN	73% ≤ ACC ≥ 96%	1000(1000)
Cobo et al., Palacio et al. (2012) [75],[76]			Texture analysis + CRF + *Mahalanobis* distance + PCA + RF	C(+)	QN	ACC = 90.27% *, SEN = 95.1%, SPE = 85.7%	103(957)
Bonet-Carne et al. (2015) [77]			Invariant features + (Regression, Trees, NN)	C(+)	QN	SEN = 86.2%, SPE = 87.0%	144(144)

Table 5
Part IV - Summary of all works (123 articles) included in this review.

Article	Modality	Organ	Method	Auto	Performance		Val
					Eval	*Value*	
Subramanian et al. (1997) [96]			RG + Gaussian / median filters	S(-)	QN	Error = 4.18%	6(6)
Piccoli et al. (1999) [44]			Texture features + (NN, SOM + K-means + Hubert index)	S(+)	QL	Visual	-
Lassige et al. (2000) [47]			Filtering + Enhancement + AC	S(-)	QN	RMSE = 2.7 px	60(240)
Siqueira et al. (2002) [45]			Statistical features + SOM + K-means	S(-)	QN	Width variation: septum = 8.78%, ventricle = 1.48% *	3(3)
Dindoyal et al. (2007a) [49]			Deformable model + SP + (Non)-Rigid registration	S(+)	QN	RMSE $\simeq$ 1.5 - 2 px	53(53)
Dindoyal et al. (2007b) [48]			Edge / Region / Shape-based level set + SP	S(+)	QN	RMSE = 2 px	26(26)
de Siqueira et al. (2007) [52]			Entropy feature + ROI isolation	S(-)	QN	Error = 11%	33(33)
Tutschek and Sahn (2008) [46]			*Semi-automatically obstetric application*	S(=)	QL	Visual	6(6)
Deng et al. (2010) [53]			Motion summation + Gaussian pyramid + Thresholding + Morphological operations + *Rayleigh-trimmed* anisotropic diffusion + AAM	S(+)	QN	Error = 8.4%	258(258)
Dindoyal et al. (2011) [51]			Unconstrained level set deformable model	S(-)	QN	Error > 10%	-
Deng et al. (2012) [54]			Motion summation + Gaussian pyramid + Thresholding + Morphological operations + *Rayleigh-trimmed* anisotropic diffusion + AAM	S(+)	QN	Error = 10%	7(7)
Sriraam et al. (2012a) [50]			Median / morphological filters + SP	S(+)	-	-	8(8)
Sriraam et al. (2012b) [56]	2D US	Heart	Morphological operations + Automated ROI	S(+)	-	-	-
Jacob et al. (2013) [57]			Median filter + K-means + AAM + Statistical features + Eigen decomposition	S(+)	QL	Visual	3(3)
Vijayalakshmi et al. (2013) [55]			Superimposed image + Connected component labelling	S(+)	QL	Visual	13(13)
Vargas-Quintero and Escalante-Ramírez (2013) [60]			Hermite transform + Wavelet decomposition + Bayesian estimator + Deformable model	-	-	-	-
Sampath and Sivaraj (2014) [63]			PPBMLE + FC	S(-)	QN	DC = 0.986, TC = 0.976 *	-
Dewi et al. (2014) [43]			*Review*	S(+/-)	QN / QL	-	-
Sardsud et al. (2015) [64]			Optical flow + C-means + Opening	S(+)	QN	Error = 2.17%	14(14)
Sridevia and Nirmala (2016) [65]			PPBMLE + FC + Texture features + FDR + ANFIS + SFIS	S(-)	QN	Error = 0.013 *, AUC = 0.8954, F-score = 0.9673, Kappa = 0.7914	224(224)
Nirmala and Sridevia (2016) [66]			PPBMLE + MRF + Texture features + RG	S(+)	QN	Error = 9.83%	4(4)
Vargas-Quintero et al. (2016a) [58]			Multitexture AAM + Hermite transform	S(+)	QN	DC = 0.8631	143(143)
Vargas-Quintero et al. (2016b) [59]			Hermite transform + Point distribution model	S(+)	-	-	-
Yu et al. (2017) [61]			Dynamic CNN	S(+)	QN	DC = 0.94	51(51)
Sundaresan et al. (2017) [62]	(videos)		Fully CNNs	S(+)	QN	Error = 23.48%	12(91)
Bridge et al. (2017) [67]	(videos)		Bayesian filtering + Intensity, gradient, motion features + Regression forests + Particle filtering	S(+)	QN	Error = 20%	12(91)
Maraci et al. (2017) [98]	(videos)		Bridge et al. (2017) + Heartbeat detection	S(+)	QN	ACC = 72.92%	323(323)
Huang et al. (2017) [95]	(videos)		CNN + Anchor mechanism + BiLSTM	S(+)	QN	Error = 27.7%	12(91)
Gao and Alison-Noble (2017) [97]	(videos)		End-to-end two-stream fully CNN	S(+)	QN	ACC = 90%	412(412)

(*) The mean value of all observers involved during the experimentation part has been calculated.
Nomenclature: Random Forest (RF), Conditional Random Field (CRF), Online Random Forest (ORF), Dynamically Balanced Online Bagging (DyBa), Convolutional Neural Network (CNN), Bidirectional Long-Short Term Memory Network (BiLSTM), Patch-to-Volume Registration (PVR), K-Nearest Neighbor (KNN), Neural Network (NN), Support Vector Machine (SVM), Histogram Equalization (HE), Region Growing (RG), MultiLinear Regression (MLR), Fast Local Descriptor for Dense Matching (DAISY), Vector Locally Aggregated Descriptor (VLAD), Gaussian Mixture Model (GMM), Principal Component Analysis (PCA), Local Intensity Order Pattern (LIOP), Fisher Vector (FV), Active Contours (AC), Gray-Level Co-occurrence Matrix (GLCM), Expectation Maximization (EM), Radial Basis Function (RBF), Self-Organizing Maps (SOM), Shape Prior (SP), Fuzzy C-Means (FCM), Probabilistic Patch Based Maximum Likelihood Estimation (PPBMLE), Fuzzy Connectedness (FC), Fisher Discriminant Ratio (FDR), Adaptive Neuro Fuzzy Inference System (ANFIS), Sugeno Fuzzy Inference System (SFIS), Markov Random Field (MRF), Active Appearance Model (AAM), Dice Coefficient (DC), Tanimoto Coefficient (TC), Root Mean Square Error (RMSE), Accuracy (ACC), Area Under Curve (AUC), Sensitivity (SEN), and Specificity (SPE).

References

Adzick, N.S., Thom, E.A., Spong, C.Y., Brock, J.W., Burrows, P.K., Johnson, M.P., Howell, L.J., Farrell, J.A., Dabrowiak, M.E., Sutton, L.N., Gupta, N., Tulipan, N.B., D'Alton, M.E., Farmer, D.L., 2011. A randomized trial of prenatal versus postnatal repair of myelomeningocele. Eng. J. Med. 364 (11), 993–1004.

Alansary, A., Kamnitsas, K., Davidson, A., Khlebnikov, R., Rajchl, M., Malamateniou, C., Rutherford, M., Hajnal, J.V., Glocker, B., Rueckert, D., Kainz, B., 2016. Fast fully automatic segmentation of the human placenta from motion corrupted MRI. In: Medical Image Computing and Computer-Assisted Intervention 2016 (MICCAI 2016), 9901, pp. 589–597.

Anquez, J., Angelini, E.D., Bloch, I., 2009. Automatic segmentation of head structures on fetal MRI. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 109–112.

Anquez, J., Angelini, E.D., Grangé, G., Bloch, I., 2013. Automatic segmentation of antenatal 3d ultrasound images. IEEE Trans. Biomed. Eng. 60 (5), 1388–1400.

Anquez, J., Bibin, L., Angelini, E.D., Bloch, I., 2010. Segmentation of the fetal envelope on ante-natal MRI. IEEE International Symposium on Biomedical Imaging: From Nano to Macro.

Baschat, A.A., 2004. Fetal responses to placental insufficiency: an update. BJOG: 111, 1031–1041.

Baumgartner, C.F., Kamnitsas, K., Matthew, J., Fletcher, T.P., Smith, S., Koch, L.M., Kainz, B., Rueckert, D., 2017. Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. IEEE Trans. Med. Imaging 36 (11), 2204–2215.

Baumgartner, C.F., Kamnitsas, K., Matthew, J., Smith, S., Kainz, B., Rueckert, D., 2016. Real-time standard scan plane detection and localisation in fetal ultrasound using fully convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9901, pp. 203–211.

Becker, B.G., Cosío, F.A., Huerta, M.E.G., Benavides-Serralde, J.A., 2010. Automatic segmentation of the cerebellum of fetuses on 3D ultrasound images, using a 3D point distribution model. In: International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4731–4734.

Benavides-Serralde, A., Hernández-Andrade, E., Fernández-Delgado, J., Plasencia, W., Scheier, M., Crispi, F., Figueras, F., Nicolaides, K.H., Gratacós, E., 2009. Three-dimensional sonographic calculation of the volume of intracranial structures in growth-restricted and appropriate-for-gestational age fetuses. Ultrasound Obstetrics Gynecol. 33, 530–537.

Benkarim, O.M., Sanroma, G., Zimmer, V.A., Muoz-Moreno, E., Hahner, N., Eixarch, E., Camara, O., González Ballester, M.A., Piella, G., 2017. Toward the automatic quantification of in utero brain development in 3d structural MRI: a review. Hum. Brain Mapp. 38 (5), 2772–2787.

Bhanu-Prakash, K.N., Ramakrishnan, A.G., Suresh, S., Chow, T.W.P., 2002. Fetal lung maturity analysis using ultrasound image features. IEEE Trans. Inf. Technol. Biomed. 6 (1), 38–45.

Bhattacharya, S., Mariani, T.J., 2013. Systems biology approaches to identify developmental bases for lung diseases. Pediatr. Res. 73, 514–522.

Bibin, L., Anquez, J., Angelini, E., Bloch, I., 2009. Hybrid 3d pregnant woman and fetus modeling from medical imaging for dosimetry studies. Int. J. Comput. Assist. Radiol. Surg. 5 (49), 1–8.

Bonet-Carne, E., Palacio, M., Cobo, T., Perez-Moreno, A., Lopez, M., Piraquive, J.P., Ramirez, J.C., Botet, F., Marques, F., Gratacós, E., 2015. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. Ultrasound Obstetrics Gynecol. 45, 427–433.

Bridge, C.P., Ioannou, C., Alison-Noble, J., 2017. Automated annotation and quantitative description of ultrasound videos of the fetal heart. Med. Image Anal. 36, 147–161.

Carneiro, G., Amat, F., Georgescu, B., Good, S., Comaniciu, D., 2008. Semantic-based indexing of fetal anatomies from 3-D ultrasound data using global / semi-local context and sequential sampling. In: IEEE Conference on Computer Vision and Pattern Recognition.

Carneiro, G., Georgescu, B., Good, S., Comaniciu, D., 2008. Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. IEEE Trans. Med. Imaging 27 (9), 1342–1355.

Ceresa, M., Torrents-Barrena, J., Alises, A., Masoller, N., Eixarch, E., Gratacós, E., González Ballester, M.A., 2017. Surgical planning system for twin-to-twin transfusion syndrome fetal surgery. In: International Conference of Computed Assisted Radiology and Surgery (CARS), 12, pp. S100–S101.

Ceresa, M., Torrents-Barrena, J., Masoller, N., Eixarch, E., González Ballester, M.A., 2017. A novel surgical planning system for TTTS. Congreso Anual de la Sociedad Espaola de Bioingeniería.

Cerrolaza, J.J., Oktay, O., Gomez, A., Matthew, J., Knight, C., Kainz, B., Rueckert, D., 2017. Fetal skull segmentation in 3D ultrasound via structured geodesic random forest. In: International Workshop on Fetal, Infant and Ophthalmic Medical Image Analysis (MICCAI), 10554, pp. 25–32.

Chen, C., Su, H., Pai, S., Hsieh, C., Jong, T., Hsu, C., Chou, S., 2011. Evaluation of placental maturity by the sonographic textures. Arch. Gynecol. Obstet. 284, 13–18.

Chen, H., Dou, Q., Ni, D., Cheng, J.-Z., Qin, J., Li, S., Heng, P.-A., 2015. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9349, pp. 507–514.

Chen, H., Ni, D., Qin, J., Li, S., Yang, X., Wang, T., Heng, P.A., 2015. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. IEEE J. Biomed. Health Inform. 19 (5), 1627–1636.

Chen, H., Ni, D., Yang, X., Li, S., Heng, P.A., 2014. Fetal abdominal standard plane localization through representation learning with knowledge transfer. In: International Workshop on Machine Learning in Medical Imaging, 8679, pp. 125–132.

Chikop, S., Shiradon, S., Poojar, P., Arun, A., Muguru-Prabhuswamy, M.S., Pungavkar, S., Geethanath, S., 2014. Segmentation and visualization of brain and lung volumes in fetal MRI using active contours and morphological operators. Int. Society for Magnetic Resonance in Medicine (ISMRM) 6606.

Choong, M.Y., Seng, M.C., Kiring, A., Yang, S.S., Teo, K.T.K., 2012. Variational level set algorithm in image segmentation for foetus ultrasound imaging system. Int. J. Simulation 13 (3), 18–23.

Cobo, T., Bonet-Carne, E., Martínez-Terrón, M., Perez-Moreno, A., Elías, N., Luque, J., Amat-Roldan, I., Palacio, M., 2012. Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. Fetal. Diagn. Ther. 31, 230–236.

Cuingnet, R., Somphone, O., Mory, B., Prevost, R., Yaqub, M., Napolitano, R., Papageorghiou, A., Roundhill, D., Noble, J., Ardon, R., 2013. Where is my baby? a fast fetal head auto-alignment in 3D-ultrasound. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 768–771.

Dahdouh, S., Andescavage, N., Yewale, S., Yarish, A., Lanham, D., Bulas, D., du Plessis, A.J., Limperopoulos, C., 2018. In vivo placental MRI shape and textural features predict fetal growth restriction and postnatal outcome. J. Magn. Reson. Imaging 47 (2), 449–458.

Dahdouh, S., Angelini, E.D., Grangé, G., Bloch, I., 2015. Segmentation of embryonic and fetal 3d ultrasound images based on pixel intensity distributions and shape priors. Med. Image Anal. 24 (1), 255–268.

Dahdouh, S., Serrurier, A., Grangé, G., Angelini, E.D., Bloch, I., 2013. Segmentation of fetal envelope from 3D ultrasound images based on pixel intensity statistical distribution and shape priors. IEEE 10^{th} International Symposium on Biomedical Imaging.

Deng, Y., Wang, Y., Chen, P., 2010. Automated detection of fetal cardiac structure from first-trimester ultrasound sequences. In: 3rd International Conference on Biomedical Engineering and Informatics, pp. 127–131.

Deng, Y., Wang, Y., Shen, Y., Chen, P., 2012. Active cardiac model and its application on structure detection from early fetal ultrasound sequences. Comput. Medical Imaging Graphics 36, 239–247.

DeSilva, M., Munoz, F.M., Mcmillan, M., Tse-Kawai, A., Marshall, H., Macartney, K.K., Joshi, J., Oneko, M., Elliott-Rose, A., Dolk, H., Trotta, F., Spiegel, H., Tomczyk, S., Shrestha, A., Kochhar, S., Kharbanda, E.O., 2016. Congenital anomalies: case definition and guidelines for data collection, analysis, and presentation of immunization safety data. Vaccine 34 (49), 6015–6026.

Dewi, D.E.O., Abduljabbar, H.N., Supriyanto, E., 2014. Review on advanced techniques in 2-d fetal echocardiography: an image processing perspective. Adv. Med. Diagnostic Technol. Chapter 3, 53–74.

Dindoyal, I., Lambrou, T., Deng, J., Todd-Pokropek, A., 2007. Automatic segmentation of low resolution fetal cardiac data using snake with shape priors. In: 5th International Symposium on image and Signal Processing and Analysis, pp. 538–543.

Dindoyal, I., Lambrou, T., Deng, J., Todd-Pokropek, A., 2007. Level set snake algorithms on the fetal heart. In: 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 864–867.

Dindoyal, I., Lambrou, T., Deng, J., Todd-Pokropek, A., 2011. 2D / 3d fetal cardiac dataset segmentation using a deformable model. Med. Phys. 38 (7), 4338–4349.

Fichter, M.A., Dornseifer, U., Henke, J., Schneider, K.T.M., Kovacs, L., Biemer, E., Bruner, J., Adzick, N.S., Harrison, M.R., Papadopulos, N.A., 2008. Fetal spina bifida repair - current trends and prospects of intrauterine neurosurgery. Fetal. Diagn. Ther. 23, 271–286.

Gao, Y., Alison-Noble, J., 2017. Detection and characterization of the fetal heartbeat in free-hand ultrasound sweeps with weakly-supervised two-streams convolutional networks. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2017), pp. 305–313.

Grannum, P., Berkowitz, A., Hobbins, J., 1979. The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity. Am. J. Obst. Gynecol. 133 (8), 915–922.

Gupta, L., Sisodia, R.S., Pallavi, V., Firtion, C., Ramachandran, G., 2011. Segmentation of 2D fetal ultrasound images by exploiting context information using conditional random fields. In: 33^{rd} Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 7219–7222.

Gutiérrez-Becker, B., Cosío, F.A., Huerta, M.E.G., Benavides-Serralde, J.A., Camargo–Marín, L., Bauelos, V.M., 2013. Automatic segmentation of the fetal cerebellum on ultrasound volumes, using a 3d statistical shape model. Med. Biol. Eng. Comput. 51, 1021–1030.

Huang, W., Bridge, C.P., Alison-Noble, J., Zisserman, A., 2017. Temporal HeartNet: Towards human-level automatic analysis of fetal cardiac screening video. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2017), pp. 341–349.

Hunter, L.E., Simpson, J., 2014. Prenatal screening for structural congenital heart disease. Nat. Rev. Cardiol. 11, 323–334.

Ison, M., Dittrich, E., Donner, R., Kasprian, G., Prayer, D., Langs, G., 2012. Fully automated brain extraction and orientation in raw fetal MRI. In: Workshop on Paediatric and Perinatal Imaging (MICCAI 2012), pp. 17–24.

Izatt, M.T., Thorpe, P.L.P.J., Thompson, R.G., D'Urso, P.S., Adam, C.J., Earwaker, J.W.S., Labrom, R.D., Askin, G.N., 2007. The use of physical biomodelling in complex spinal surgery. Eur. Spine J. 16, 1507–1518.

Kainz, B., Keraudren, K., Kyriakopoulou, V., Rutherford, M., Hajnal, J.V., Rueckert, D., 2014. Fast fully automatic brain detection in fetal MRI using dense rotation invariant image descriptors. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1230–1233.

Kainz, B., Malamateniou, C., Murgasova, M., Keraudren, K., Rutherford, M., Hajnal, J.V., Rueckert, D., 2014. Motion corrected 3D reconstruction of the fetal thorax from prenatal MRI. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI 2014), 8674, pp. 284–291.

Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78.

Keraudren, K., Kainz, B., Oktay, O., Kyriakopoulou, V., Rutherford, M., Hajnal, J.V., Rueckert, D., 2015. Automated localization of fetal organs in MRI using random forests with steerable features. In: Medical Image Computing and Computer-Assisted Intervention 2015 (MICCAI 2015), 9351, pp. 620–627.

Keraudren, K., Kuklisova-Murgasova, M., Kyriakopoulou, V., Malamateniou, C., Rutherford, M., Kainz, B., Hajnal, J., Rueckert, D., 2014. Automated fetal brain segmentation from 2d MRI slices for motion correction. Neuroimage 101, 633–643.

Khalili, N., Moeskops, P., Claessens, N.H.P., Scherpenzeel, S., Turk, E., de Heus, R., Benders, M.J.N.L., Viergever, M.A., Pluim, J.P.W., Išgum, I., 2017. Automatic segmentation of the intracranial volume in fetal MR images. In: International Workshop on Fetal, Infant and Ophthalmic Medical Image Analysis, 10554, pp. 42–51.

Kohl, T., Kawecki, A., Degenhardt, J., Axt-Fliedner, R., 2012. Preoperative sonoanatomic examination of fetal spina bifida aperta permits prediction of surgical complexity during subsequent minimally-invasive fetoscopic closure. In: 22nd World Congress on Ultrasound in Obstetrics and Gynecology, 40, pp. 1–54.

Kording, F., Forkert, N.D., Sedlacik, J., Adam, G., Hecher, K., Arck, P., Remus, C., 2015. Automatic differentiation of placental perfusion compartments by time-to-peak analysis in mice. Placenta 36, 255–261.

Krishnan, P.H., Karthickeyan, V., Ramamoorthy, P., 2014. A novel method for measurement of fetal volume from US images using segmentation techniques. In: International Conference on Green Computing Communication and Electrical Engineering.

Lassige, T.A., Benkeser, P.J., Fyfe, D., Sharma, S., 2000. Comparison of septal defects in 2d and 3d echocardiography using active contour models. Comput. Med. ImagingGraph. 24, 377–388.

Lean, S.C., Heazell, A.E.P., Dilworth, M.R., Mills, T.A., Jones, R.L., 2017. Placental dysfunction underlies increased risk of fetal growth restriction and stillbirth in advanced maternal age women. Sci Rep 7 (9677), 1–16.

Lei, B., Li, X., Yao, Y., Li, S., Chen, S., Zhou, Y., Ni, D., Wang, T., 2014. Automatic grading of placental maturity based on LIOP and Fisher Vector. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4671–4674.

Lei, B., Tan, E., Chen, S., Zhuo, L., Li, S., Ni, D., Wang, T., 2015. Automatic recognition of fetal facial standard plane in ultrasound image via fisher vector. PLoS ONE 1–20.

Lei, B., Yao, Y., Chen, S., Li, S., Li, W., Ni, D., Wang, T., 2015. Discriminative learning for automatic staging of placental maturity via multi-layer fisher vector. Sci Rep 5 (12818), 1–11.

LemosDeSiqueira, M., Muller, D.N., Navaux, P.O.A., 2007. Cardiac structure recognition in ultrasound images. In: 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services, pp. 463–466.

LemosDeSiqueira, M., Scharcanski, J., Navaux, P.O.A., 2002. Echocardiographic image sequence segmentation and analysis using self-organizing maps. J. VLSI Sig. Proc. 32, 135–145.

Li, X., Wang, Y., Yu, J., Chen, P., 2010. Fetal lung segmentation using texture-based boundary enhancement and active contour models. In: 3rd International Conference on Biomedical Engineering and Informatics (BMEI), pp. 264–268.

Li, X., Yao, Y., Ni, D., Chen, S., Li, S., Lei, B., Wang, T., 2014. Automatic staging of placental maturity based on dense descriptor. Biomed. Mater. Eng. 24, 2821–2829.

Li, Y., Cerrolaza, J.J., Sinclair, M., Hou, B., Alansary, A., Khanal, B., Matthew, J., Kainz, B., Rueckert, D., 2018. Standard plane localisation in 3D fetal ultrasound using network with geometric and image loss. In: 1st Conference on Medical Imaging with Deep Learning (MIDL).

Li, Y., Khanal, B., Hou, B., Alansary, A., Cerrolaza, J.J., Sinclair, M., Matthew, J., Gupta, C., Knight, C., Kainz, B., Rueckert, D., 2018. Standard plane detection in 3D fetal ultrasound using an iterative transformation network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).

Li, Y., Xu, R., Ohya, J., Iwata, H., 2017. Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers. In: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Liao, H., Tsuzuki, M., Mochizuki, T., Kobayashi, E., Chiba, T., Sakuma, I., 2009. Fast image mapping of endoscopic image mosaics with three-dimensional ultrasound image for intrauterine fetal surgery. Minimally Invasive Therapy Allied Technol. 18 (6), 332–340.

Linares, P.A., McCullagh, P.J., Black, N.D., Dornan, J., 2003. Characterization of ultrasonic images of the placenta based on textural features. In: 4th International IEEE Conference on Information Technology Applications in Biomedicine, pp. 211–214.

Linares, P.A., McCullagh, P.J., Black, N.D., Dornan, J., 2004. Feature selection for the characterization of ultrasonic images of the placenta using texture classification. In: IEEE International Symposium on Biomedical Imaging: Nano to Macro, pp. 1147–1150.

Liu, X., Annangi, P., Gupta, M., Yu, B., Padfield, D., Banerjee, J., Krishnan, K., 2012. Learning-based scan plane identification from fetal head ultrasound images. SPIE - The International Society for Optical Engineering, 8320.

Liu, Z., Zheng, H., Lin, S., 2009. Application of multi-classification support vector machine in the B-placenta image classification. In: International Conference on Computational Intelligence and Software Engineering, pp. 1–4.

Looney, P., Stevenson, G.N., Nicolaides, K.H., Plasencia, W., Molloholli, M., Natsis, S., Collins, S.L., 2017. Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning. In: IEEE 14th International Symposium on Biomedical Imaging 2017 (ISBI 2017), pp. 279–282.

Looney, P., Stevenson, G.N., Nicolaides, K.H., Plasencia, W., Molloholli, M., Natsis, S., Collins, S.L., 2018. Fully automated, real-time 3d ultrasound segmentation to estimate first trimester placental volume using deep learning. JCI Insight 3 (11), 1–9.

Luks, F.I., Carr, S.R., Ponte, B., Rogg, J.M., T. F. T., 2001. Preoperative planning with magnetic resonance imaging and computerized volume rendering in twin–to-twin transfusion syndrome. Am. J. Obstetrics Gynecol. 185 (1), 216–219.

Luo, J., Turk, E.A., Bibbo, C., Gagoski, B., Roberts, D.J., Vangel, M., Tempany-Afdhal, C.M., Barnewolt, C., Estroff, J., Palanisamy, A., Barth, W.H., Zera, C., Malpica, N., Golland, P., Adalsteinsson, E., Robinson, J.N., Grant, P.E., 2017. In vivo quantification of placental insufficiency by BOLD MRI: a human study. Sci. Rep. 7 (3713), 1–10.

Makropoulos, A., Counsell, S.J., Rueckert, D., 2018. A review on automatic fetal and neonatal brain MRI segmentation. Neuroimage 170, 231–248.

Malathi, G., Shanthi, V., 2009. Wavelet based features for ultrasound placenta images classification. In: Second International Conference on Emerging Trends in Engineering and Technology (ICETET-09), pp. 341–345.

Malathi, G., Shanthi, V., 2010. Histogram based classification of ultrasound images of placenta. Int. J. Comput. Appl. 1 (16), 49–53.

Malathi, G., Shanthi, V., 2010. Statistical measurement of ultrasound placenta images using segmentation approach. In: International Conference on Signal and Image Processing, pp. 309–316.

Malathi, G., Shanthi, V., 2010. Thickness based characterization of ultrasound placenta images using regression analysis. Int. J. Comput. Appl. 3 (7), 7–11.

Malathi, G., Shanthi, V., 2011. Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach. J. Inf. Hiding Multimedia Signal Process. 2 (4), 332–343.

Maraci, M.A., Bridge, C.P., Napolitano, R., Papageorghiou, A., Alison-Noble, J., 2017. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. Med. Image Anal. 37, 22–36.

Maselli, K.M., Badillo, A., 2016. Advances in fetal surgery. Ann. Transl. Med. 4 (20:394), 1–7.

Mathis, J., Raio, L., Baud, D., 2015. Fetal laser therapy: applications in the management of fetal pathologies. Prenat. Diagn. 35, 623–636.

Matturri, L., Lavezzi, A.M., Minoli, I., Ottaviani, G., Rubino, B., Cappellini, A., Rossi, L., 2003. Association between pulmonary hypoplasia and hypoplasia of arcuate nucleus in stillbirth. J. Perinatol. 23, 328–332.

Melbourne, A., Pratt, R., Owen, D., Sokolska, M., Bainbridge, A., Atkinson, D., Kendall, G., Deprest, J., Vercauteren, T., David, A., Ourselin, S., 2016. Placental image analysis using coupled diffusion-weighted and multi-echo T2 MRI and a multi-compartment model. International Workshop on Perinatal, Preterm and Paediatric Image analysis (MICCAI).

Mesas-Burgos, C., Frenckner, B., Luco, M., Harting, M.T., Lally, P.A., Lally, K.P., 2017. Prenatally diagnosed congenital diaphragmatic hernia: optimal mode of delivery? J. Perinatol. 37, 134–138.

Miao, H., Mistelbauer, G., Karimov, A., Alansary, A., Davidson, A., Lloyd, D.F.A., Damodaram, M., Story, L., Hutter, J., Hajnal, J.V., Rutherford, M., Preim, B., Kainz, B., Groller, M.E., 2017. Placenta maps: in utero placental health assessment of the human fetus. IEEE Trans. Vis. Comput Graph 23 (6), 1612–1623.

Moore, R.J., Strachan, B.K., Tyler, D.J., Duncan, K.R., Baker, P.N., Worthington, B.S., Johnson, I.R., Gowland, P.A., 2000. In utero perfusing fraction maps in normal and growth restricted pregnancy measured using IVIM echo-planar MRI. Placenta 21 (7), 726–732.

Moran, M., Higgins, M., Zombori, G., Ryan, J., McAuliffe, F.M., 2013. Computerized assessment of placental calcification post-ultrasound: a novel software tool. Ultrasound Obstetrics Gynecol. 41, 545–549.

Moran, M., Ryan, J., Brennan, P.C., Higgins, M., McAuliffe, F.M., 2009. Is grannum grading of the placenta reproducible? Proc. SPIE Int. Society Optical Eng. 7263, 72631W–1–72631W–6.

Napolitano, R., Yaqub, M., Ioannou, C., Knight, C.L., Noble, J.A., Papageorghiou, A.T., 2012. Automatic detection of fetal brain structures from ultrasound volumes. In: 22nd World Congress on Ultrasound in Obstetrics and Gynecology, 40, p. 134.

Ni, D., Li, T., Yang, X., Qin, J., Li, S., Chin, C.-T., Ouyang, S., Wang, T., Chen, S., 2013. Selective search and sequential detection for standard plane localization in ultrasound. In: International Workshop on Computational and Clinical Challenges in Abdominal Imaging, 8198, pp. 203–211.

Ni, D., Yang, X., Chen, X., Chin, C.-T., Chen, S., Heng, P.A., Li, S., Qin, J., Wang, T., 2014. Standard plane localization in ultrasound by radial component model and selective search. Ultrasound Med. Biol. 40 (11), 2728–2742.

Nirmala, S., Sridevi, S., 2016. Markov random field segmentation based sonographic identification of prenatal ventricular septal defect. In: Procedia Computer Science - 7th International Conference on Communication, Computing and Virtualization, 79, pp. 344–350.

Oguz, I., Pouch, A., Yushkevich, N., Wang, H., Gee, J., Schwartz, N., Yushkevich, P., 2016. Fully automated placenta segmentation from 3D ultrasound images. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2016).

Palacio, M., Cobo, T., Martínez-Terrón, M., Rattá, G.A., Bonet-Carné, E., Amat-Roldán, I., Gratacós, E., 2012. Performance of an automatic quantitative ultrasound analysis of the fetal lung to predict fetal lung maturity. Am. J. Obstetrics Gynecol. 504, 1–5.

Penney, G.P., Blackall, J.M., Hamady, M.S., Sabharwal, T., Adam, A., Hawkes, D.J., 2004. Registration of freehand 3d ultrasound and magnetic resonance liver images. Med. Image Anal. 8, 81–91.

Phillips, T.J., Scott, H., Menassa, D.A., Bignell, A.L., Sood, A., Morton, J.S., Akagi, T., Azuma, K., Rogers, M.F., Gilmore, C.E., Inman, G.J., Grant, S., Chung, Y., Aljunaidy, M.M., Cooke, C.-L., Steinkraus, B.R., Pocklington, A., Logan, A., Collett, G.P., Kemp, H., Holmans, P.A., Murphy, M.P., Fulga, T.A., Coney, A.M., Akashi, M., Davidge, S.T., Case, C.P., 2017. Treating the placenta to prevent adverse effects of gestational hypoxia on fetal brain development. Sci. Rep. 7 (9079), 1–16.

Piccoli, L., Dahmer, A., Scharcanski, J., Navaux, P.O.A., 1999. Fetal echocardiographic image segmentation using neural networks. In: 7th International Conference on Image Processing and its Applications, 465, pp. 507–511.

Postma, A.V., Bezzina, C., Christoffels, V.M., 2016. Genetics of congenital heart disease: the contribution of the noncoding regulatory genome. J. Hum. Genet. 61, 13–19.

Pratt, R., Deprest, J., Vercauteren, T., Ourselin, S., David, A.L., 2015. Computer-assisted surgical planning and intraoperative guidance in fetal surgery: a systematic review. Prenat. Diagn. 35 (12), 1159–1166.

Rahmatullah, B., Papageorghiou, A., Noble, J.A., 2011. Automated selection of standardized planes from ultrasound volume. In: International conference on Machine learning in medical imaging, 7009, pp. 35–42.

Rajchl, M., Lee, M., Oktay, O., Kamnitsas, K., Passerat-Palmbach, J., Bai, W., Rutherford, M., Hajnal, J., Kainz, B., Rueckert, D., 2017. Deepcut: object segmentation from bounding box annotations using convolutional neural networks. IEEE Trans. Med. Imaging 36 (2), 674–683.

Reddy, U.M., Filly, R.A., Copel, J.A., 2008. Prenatal imaging: ultrasonography and magnetic resonance imaging. Obstetrics Gynecol. 112 (1), 145–157.

Reshmi-Mariam-Reji-Jacob, Prabakar, S., Porkumaran, K., 2013. Fetal cardiac structure detection from ultrasound sequences. Int. J. Instrum. Control and Autom. (IJICA) 2 (1), 12–16.

Roy-Lacroix, M.E., Moretti, F., Ferraro, Z.M., Brosseau, L., Clancy, J., Fung-Kee-Fung, K., 2017. A comparison of standard two-dimensional ultrasound to three-dimensional volume sonography for routine second-trimester fetal imaging. J. Perinatol. 37, 380–386.

Roy-Lacroix, M.E., Moretti, F., Ferraro, Z.M., Brosseau, L., Clancy, J., Fung-Kee-Fung, K., 2017. A comparison of standard two-dimensional ultrasound to three-dimensional volume sonography for routine second-trimester fetal imaging. J. Perinatol. 37, 380–386.

Ryan, J.T., McAuliffe, F., Higgins, M., Stanton, M., Brennan, P., 2008. A novel software-based technique for quantifying placental calcifications and infarctions from ultrasound. Proc. SPIE Medical Imaging 2008: Ultrasonic Imaging Signal Process. 6920, 1–8.

Salehi, S.S.M., Hashemi, S.R., Velasco-Annis, C., Ouaalam, A., Estroff, J.A., Erdogmus, D., Warfield, S.K., Gholipour, A., 2018. Real-time automatic fetal brain extraction in fetal brain MRI by deep learning. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 720–724.

Sampath, S., Sivaraj, N., 2014. Fuzzy connectedness based segmentation of fetal heart from clinical ultrasound images. Adv. Comput., Networking Inf. 1, 329–337.

Sardsud, C., Auephanwiriyakul, S., Theera-Umpon, N., Tongsong, T., 2015. Patch-based fetal heart chamber segmentation in ultrasound sequences using possibilistic clustering. In: 7th International Conference on Computational Intelligence, Modelling and Simulation, pp. 32–37.

Schlemper, J., Oktay, O., Chen, L., Matthew, J., Knight, C., Kainz, B., Glocker, B., Rueckert, D., 2018. Attention-gated networks for improving ultrasound scan plane detection. In: 1st Conference on Medical Imaging with Deep Learning (MIDL).

Schneider, R.J., Perrin, D.P., Vasilyev, N.V., Marx, G.R., del Nido, P.J., Howe, R.D., 2012. Mitral annulus segmentation from four-dimensional ultrasound using a valve state predictor and constrained optical flow. Med. Image Anal. 16 (2), 497–504.

Schneider, R.J., Tenenholtz, N.A., Perrin, D.P., Marx, G.R., del Nido, P.J., Howe, R.D., 2011. Patient-specific mitral leaflet segmentation from 4D ultrasound. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 14, pp. 520–527.

Slaghekke, F., Lopriore, E., Lewi, L., Middeldorp, J.M., van Zwet, E.W., Weingertner, A., Klumper, F.J., DeKoninck, P., Devlieger, R., Kilby, M.D., Rustico, M.A., Deprest, J., Favre, R., Oepkes, D., 2014. Fetoscopic laser coagulation of the vascular equator versus selective coagulation for twin-to-twin transfusion syndrome: an open-label randomised controlled trial. Lancet 383 (9935), 2144–2151.

Sridevia, S., Nirmala, S., 2016. ANFIS Based decision support system for prenatal detection of truncus arteriosus congenital heart defect. Appl. Soft Comput. 46, 577–587.

Sriraam, N., Vijayalakshmi, S., Suresh, S., 2012. Automated screening of fetal heart chambers from 2-d ultrasound cine-loop sequences. Int. J. Biomed. Clinical Eng. 1 (2), 24–33.

Sriraam, N., Vijayalakshmi, S., Suresh, S., 2012. Computer-aided fetal cardiac scanning using 2d ultrasound: perspectives of fetal heart biometry. Int. J. Biomed. Clinical Eng. 1 (1), 1–13.

Stevenson, G.N., Collins, S.L., Ding, J., Impey, L., Alison-Noble, J., 2015. 3-D ultrasound segmentation of the placenta using the random walker algorithm: reliability and agreement. Ultrasound Med. Biol. 41 (12), 3182–3193.

Stevenson, G.N., Collins, S.L., Impey, L., Alison-Noble, J., 2010. A novel semi-automated (SA) technique for 3D ultrasound measurement of placental volume. In: 20th World Congress on Ultrasound in Obstetrics and Gynecology, 36, p. 82.

Story, L., Rutherford, M., 2015. Advances and applications in fetal magnetic resonance imaging. Obst. Gynaecol. 17, 189–199.

Studholme, C., 2011. Mapping fetal brain development in utero using MRI: the big bang of brain mapping. Annu. Rev. Biomed. Eng. 13, 345–368.

Studholme, C., Rousseau, F., 2014. Quantifying and modelling tissue maturation in the living human fetal brain. Int. J. Dev. Neurosci. 32 (0), 3–10.

Subramanian, K.R., Lawrence, D.M., Mostafavi, M.T., 1997. Interactive segmentation and analysis of fetal ultrasound images. In: Visualization in Scientific Computing '97, pp. 115–123.

Sundaresan, V., Bridge, C.P., Ioannou, C., Alison-Noble, J., 2017. Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks. In: IEEE 14th International Symposium on Biomedical Imaging 2017 (ISBI 2017), pp. 671–674.

Taimouri, V., Gholipour, A., Velasco-Annis, C., Estroff, J.A., Warfield, S.K., 2015. A template-to-slice block matching approach for automatic localization of brain in fetal MRI. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 144–147.

Terui, K., Nagata, K., Kanamori, Y., Takahashi, S., Hayakawa, M., Okuyama, H., Inamura, N., Yoshida, H., Taguchi, T., Usui, N., 2017. Risk stratification for congenital diaphragmatic hernia by factors within 24h after birth. J. Perinatol. 0, 1–4.

Torrents-Barrena, J., Piella, G., Masoller, N., Gratacós, E., Eixarch, E., Ceresa, M., González Ballester, M.A., 2018. LSTM fully convolutional neural networks for TTTS umbilical cord segmentation. In: International Conference of Computed Assisted Radiology and Surgery (CARS), 13.

Tourbier, S., Hagmann, P., Cagneaux, M., Guibaud, L., Gorthi, S., Schaer, M., Thiran, J.-P., Meuli, R., Cuadra, M.B., 2015. Automatic brain extraction in fetal MRI using multi-atlas-based segmentation. SPIE Medical Imaging 2015: Image Processing 94130Y.

Tutschek, B., Sahn, D.J., 2008. Semi-automatic segmentation of fetal cardiac cavities: progress towards an automated fetal echocardiogram. Ultrasound Obstetrics Gynecol. 32, 176–180.

Valsky, D.V., Eixarch, E., Martinez-Crespo, J.M., Acosta, E.-R., Lewi, L., Deprest, J., Gratacós, E., 2011. Fetoscopic laser surgery for twin-to-twin transfusion syndrome after 26 weeks of gestation. Fetal. Diagn. Ther. 31 (1), 30–34.

Vargas-Quintero, L., Escalante-Ramírez, B., 2013. Filtering and left ventricle segmentation of the fetal heart in ultrasound images. In: Proc. SPIE 8922, IX International Seminar on Medical Information Processing and Analysis, 89220X.

Vargas-Quintero, L., Escalante-Ramírez, B., Camargo-Marín, L., Guzmán-Huerta, M., Arámbula-Cosio, F., Borboa-Olivares, H., 2016. Left ventricle segmentation in fetal echocardiography using a multi-texture active appearance model based on the steered hermite transform. Comput. Methods Programs Biomed. 137, 231–245.

Vargas-Quintero, L., Escalante-Ramírez, B., Camargo-Marín, L., Guzmán-Huerta, M., Arámbula-Cosio, F., Borboa-Olivares, H., 2016. Shape extraction in fetal ultrasound images using a hermite-based filtering approach and a point distribution model. Proc. SPIE 9896, Optics, Photonics and Digital Technologies for Imaging Applications IV 98961G.

Velásquez-Rodríguez, G., Cosío, F.A., Huerta, M.E.G., Marín, L.C., Olivares, H.B., Ramírez, B.E., 2015. Automatic segmentation of the cerebellum in ultrasound volumes of the fetal brain. Revista Mexicana de Ingeniería Biomédica 36 (2), 121–129.

Velásquez-Rodríguez, G., Cosío, F.A., Ramírez, B.E., 2015. Automatic segmentation of the fetal cerebellum using spherical harmonics and gray level profiles. International Symposium on Medical Information Processing and Analysis, 9681.

Vijayalakshmi, S., Sriraam, N., Suresh, S., Muttan, S., 2013. Automated region mask for four-chamber fetal heart biometry. J. Clin. Monit. Comput. 27 (2), 205–209.

Wang, G., Li, W., Zuluaga, M.A., Pratt, R., Patel, P.A., Aertsen, M., Doel, T., David, A.L., Deprest, J., Ourselin, S., Vercauteren, T., 2018. Interactive medical image segmentation using deep learning with image-specific fine-tuning. IEEE Trans. Med Imaging.

Wang, G., Zuluaga, M.A., Li, W., Pratt, R., Patel, P.A., Aertsen, M., Doel, T., David, A.L., Deprest, J., Ourselin, S., Vercauteren, T., 2018. DeepIGeos: a deep interactive geodesic framework for medical image segmentation. IEEE Trans. Med Imaging.

Wang, G., Zuluaga, M.A., Pratt, R., Aertsen, M., David, A.L., Deprest, J., Vercauteren, T., Ourselin, S., 2015. Minimally interactive placenta segmentation from motion corrupted MRI for fetal surgical planning. Workshop on Interactive Medical Image Computing (IMIC).

Wang, G., Zuluaga, M.A., Pratt, R., Aertsen, M., David, A.L., Deprest, J., Vercauteren, T., Ourselin, S., 2015. Slic-seg: Slice-by-slice segmentation propagation of the placenta in fetal MRI using one-plane scribbles and online learning. In: Medical Image Computing and Computer-Assisted Intervention 2015 (MICCAI 2015), 9351, pp. 29–37.

Wang, G., Zuluaga, M.A., Pratt, R., Aertsen, M., Doel, T., Klusmann, M., David, A.L., Deprest, J., Vercauteren, T., Ourselin, S., 2016. Dynamically balanced online random forests for interactive scribble-based segmentation. In: Medical Image Computing and Computer-Assisted Intervention 2016 (MICCAI 2016), 9901, pp. 352–360.

Wang, G., Zuluaga, M.A., Pratt, R., Aertsen, M., Doel, T., Klusmann, M., David, A.L., Deprest, J., Vercauteren, T., Ourselin, S., 2016. Slic-seg: a minimally interactive segmentation of the placenta from sparse and motion-corrupted fetal MRI in multiple views. Med. Image Anal. 34, 137–147.

Wright, R., Kyriakopoulou, V., Ledig, C., Rutherford, M., Hajnal, J., Rueckert, D., Aljabar, P., 2014. Automatic quantification of normal cortical folding patterns from fetal brain MRI. Neuroimage 91, 21–32.

Wu, C., Chen, C., Chou, H., 2017. Pulmonary hypoplasia induced by oligohydramnios: findings from animal models and a population-based study. Pediatrics Neonatol. 58 (1), 3–7.

Yan, W., 2017. Womb zoom: what advances in fetal and newborn imaging have revealed. Nat. Med. 23 (3), 270–271.

Yang, X., Ni, D., Qin, J., Li, S., Wang, T., Chen, S., Heng, P.A., 2014. Standard plane localization in ultrasound by radial component. In: International Symposium on Biomedical Imaging (ISBI), pp. 1180–1183.

Yang, X., Yu, L., Li, S., Wang, X., Wang, N., Qin, J., Ni, D., Heng, P., 2017. Towards automatic semantic segmentation in volumetric ultrasound. In: Medical Image Computing and Computer-Assisted Intervention 2017 (MICCAI 2017), 10433, pp. 711–719.

Yang, X., Yu, L., Li, S., Wang, X., Wang, N., Qin, J., Ni, D., Heng, P.-A., 2017. Towards automatic semantic segmentation in volumetric ultrasound. In: Medical Image Computing and Computer Assisted Intervention (MICCAI 2017), pp. 711–719.

Yaqub, M., Cuingnet, R., Napolitano, R., Roundhill, D., Papageorghiou, A., Ardon, R., Noble, J.A., 2013. Volumetric segmentation of key fetal brain structures in 3D ultrasound. In: International Workshop on Machine Learning in Medical Imaging, 8184, pp. 25–32.

Yaqub, M., Kelly, B., Papageorghiou, A., Noble, J.A., 2015. Guided random forests for identification of key fetal anatomy and image categorization in ultrasound scans. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9351, pp. 687–694.

Yaqub, M., Kopuri, A., Rueda, S., Sullivan, P.B., McCormick, K., Noble, J.A., 2014. A constrained regression forests solution to 3D fetal ultrasound plane localization for longitudinal analysis of brain growth and maturation. In: International conference on Machine learning in medical imaging, 8679, pp. 109–116.

Yaqub, M., Napolitano, R., Ioannou, C., Papageorghiou, A.T., Noble, J.A., 2012. Automatic detection of local fetal brain structures in ultrasound images. In: IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1555–1558.

Yoon, H.M., Kim, E.A.-R., Chung, S.-H., Kim, S.-O., Jung, A.Y., Cho, Y.A., Yoon, C.H., Lee, J.S., 2017. Extralobar pulmonary sequestration in neonates: the natural course and predictive factors associated with spontaneous regression. Eur. Radiol. 27 (6), 2489–2496.

You, W., Andescavage, N., Zun, Z., Limperopoulos, C., 2017. Semi-automatic segmentation of the placenta into fetal and maternal compartments using intravoxel incoherent motion MRI. In: Proc. SPIE 10137 - Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging.

Yu, L., Guo, Y., Wang, Y., Yu, J., Chen, P., 2017. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks. IEEE Trans. Biomed. Eng. 64 (8), 1886–1895.

Yu, Z., Tan, E., Ni, D., Qin, J., Chen, S., Li, S., Lei, B., Wang, T., 2018. A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition. IEEE J. Biomed. Health Inform. 22 (3), 874–885.

Zhang, L., Chen, S., Chin, C.T., Wang, T., Li, S., 2012. Intelligent scanning: automated standard plane selection and biometric measurement of early gestational sac in routine ultrasound examination. Med Phys. 39 (8), 5015–5027.

Zhang, T., Matthew, J., Lohezic, M., Davidson, A., Rutherford, M., Rueckert, D., Hajnal, J.V., Aljabar, P., 2016. Graph-based whole body segmentation in fetal MR images. Medical Image Computing and Computer-Assisted Intervention (MICCAI 2016).

Zito, L.L., Kadji, C., Cannie, M., Kacem, Y., Strizek, B., Mbonyumutwa, M., Wuyts, F., Jani, J., 2013. Determination of fetal body volume measurement at term with magnetic resonance imaging: effect of various factors. J. Maternal-Fetal Neonatal Med. 26 (12), 1254–1258.