Review

Machine learning in fetal cardiology: what to expect.

Patricia Garcia-Canadilla[1,2*], Sergio Sanchez-Martinez[1], Fatima Crispi[1,3] and Bart Bijnens[1,4,5]

[1] Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain
[2] Institute of Cardiovascular Science, University College London, London, UK
[3] Fetal Medicine Research Center, BCNatal - Barcelona Center for Maternal-Fetal and Neonatal Medicine (Hospital Clínic and Hospital Sant Joan de Déu), Institut Clínic de Ginecologia Obstetricia i Neonatologia, Centre for Biomedical Research on Rare Diseases (CIBER-ER), Barcelona, Spain.
[4] Department of Cardiovascular Sciences, KU Leuven, Leuven, Belgium
[5] ICREA, Barcelona, Spain

Short title: Machine learning in fetal cardiology.

***Corresponding Author**
Dr Patricia Garcia-Canadilla
Institut d'Investigacions Biomèdiques August Pi i Sunyer
C/ Rosselló, 149-153
08036 Barcelona
Spain
e-mail: patricia.garciac@upf.edu

ABSTRACT

In Fetal Cardiology, imaging, and especially echocardiography, has demonstrated to help in the diagnosis and monitoring of fetuses with a compromised cardiovascular system potentially associated to several fetal conditions. Different ultrasound approaches are currently used to evaluate fetal cardiac structure and function, including conventional 2D imaging, M-mode and Tissue Doppler Imaging among others. However, assessing the fetal heart is still challenging mainly due to involuntary movements of the fetus, the small size of the heart and the lack of expertise in fetal echocardiography of some sonographers. Therefore, the use of new technologies to improve the primary acquired images, help extracting measurements, or to aid in the diagnosis of cardiac abnormalities, is of great importance for optimal assessment of the fetal heart. Machine leaning (ML) is a computer science discipline focused on teaching a computer to perform tasks with specific goals without explicitly programming the rules on how to perform this task. In this review we provide a brief overview on the potential of ML techniques to improve the evaluation of fetal cardiac function by optimising the image acquisition and quantification/segmentation, as well as aid in improving the prenatal diagnoses of fetal cardiac remodelling and abnormalities.

Keywords: machine learning; deep learning; artificial intelligence; fetal cardiology; obstetrics; echocardiography; decision support systems.

1. Introduction

Fetal echocardiography was introduced to assess fetal cardiac function only 15 years ago (in 2004 the first study was performed). It has evolved from the description of cardiac anatomical abnormalities towards the quantitative assessment of cardiac dimensions, shape and function, and demonstrated to be useful in the diagnosis and monitoring of fetuses with a compromised cardiovascular system related to several fetal conditions, such as intrauterine growth restriction (IUGR), twin-to-twin transfusion syndrome, congenital heart disease, etc [1–3]. Moreover, some cardiac parameters have already shown to be helpful to predict perinatal problems and long-term cardiovascular outcome [4].

Different ultrasound (US) approaches are currently used to evaluate fetal cardiac function, including conventional 2D imaging, M-mode, blood-pool and Tissue Doppler Imaging (TDI), 2D speckle tracking and 4D-spatio-temporal imaging correlation (STIC) [4,5]. For any evaluation, an optimal image of the fetal heart is crucial to adequately assess cardiac structure and function. However, assessing fetal cardiac function is still challenging due to involuntary movements of the fetus, the small size of the heart, high heart rate, the limited access to the fetus and the lack of expertise in fetal echocardiography of some sonographers. After having obtained an optimal image, measurements have to be performed in order to extract relevant cardiac features that relate to remodelling and functional status. Currently, these are mainly carried out manually by the sonographer, either during the investigation or offline using a dedicated workstation. Therefore, the use of new technologies to improve the primary acquired images or help extracting and standardising measurements is of great importance for optimal assessment of the fetal heart.

Machine leaning (ML) is a computer science discipline focused on teaching a computer to perform tasks, with a specific goal in mind, without explicitly programming the rules on how to perform this task. Mathematically speaking, learning occurs when a computer iteratively improves its performance on the given task (e.g., classification of a disease or estimation of clinical measurements) with experience, or in other words, when it is exposed to data [6]. Usually, ML algorithms are classified in two approaches: supervised and unsupervised learning algorithms (see Figure 1). Deep learning (DL), a popular algorithm (and often thought of when the term Machine Learning is used) just a subset of machine learning that uses a layered structure of calculations know as artificial neural networks (ANN) on unstructured data. Figure 2 illustrates the typical pipeline for both supervised and unsupervised learning algorithms. Supervised learning requires explicit ground truth goals (diagnostic labels, outcomes, reference image measurements, etc.) from which the algorithm can optimize its performance during training. Supervised learning algorithms can be further classified into classification and regression (see Figure 1). Classification techniques evaluate the given input and come up with a category such as 'red' or 'blue' or 'disease' or 'non disease, while regression techniques result in a continuous output: the value of the predicted quantity (such as a probability of a diagnosis). Besides DL, the most common classification algorithms include decision trees, support vector machine (SVM), etc., while linear and logistic regressions are typical regression algorithms (see Figure 1). On the other hand, unsupervised learning algorithms receive unlabelled examples and aim at discovering main patterns or similarities in the data, which would correspond to different disease manifestations or different phenotypes within a given disease, or different temporal evolution. Consequently, supervised learning is commonly used when the final goal is well known at the time of learning and unsupervised learning is used as an exploratory tool and usually the final goal follows from the analysis of the obtained results. Unsupervised learning algorithms can be further classified into clustering and dimensionality reduction as illustrated in Figure

1. Typical clustering algorithms include K-means or Gaussian mixture models, while principal component analysis (PCA), and linear discriminant analysis (LCA) are classical dimensionality reduction techniques.

Once ML models are trained, their performance on unseen data (referred to as *test set*) is known as the model's generalizability (Figure 2). Models that perform considerably better on the training set compared to the *test set* are overfitted, which means that they have a big adherence to the training cases, but new patients are not correctly handled. Finding a good balance between training and testing performance is thus crucial for the application of ML models in clinical settings. A related highly relevant risk when using ML for clinical decision making is how to deal with, and not miss, rare occurrences in the (testing) data that were underrepresented in the *training dataset*. To circumvent this risk, ML approaches (and especially supervised once) need to be trained with a dataset that sufficiently captures the phenomenon under study. For clinical decision making, an unsupervised approach that highlights these rare instances might therefore be better as compared to a supervised one that forces decisions towards what was trained for. In order to learn more about ML concepts, we refer the reader to the review paper by Deo [7].

ML techniques can help to optimise image acquisition protocols, thus reducing the acquisition time and ensuring optimal quality, and can help extracting comprehensive and standardised information for a better evaluation of cardiac function. In this review we provide a brief overview on machine/deep learning applications in obstetrics with a particular focus on the evaluation of fetal cardiac function by optimising the image acquisition and quantification/segmentation, as well as aid in improving the prenatal diagnoses of fetal cardiac remodelling and abnormalities.

2. Machine learning for data acquisition

Image acquisition is the first step towards building a system to optimise the characterisation of fetal cardiac function. This step is of capital importance, as the extracted information will be greatly conditioned by the intrinsic quality and amount of input data. The acquisition of the best standard fetal views is labour-intensive and relies on the sonographer's experience. The resulting inter-operator variability in image acquisition hampers individual temporal follow-up, or the combination of different data sources for research purposes. In this sense, ML-powered acquisition methods to speed up the acquisition, decrease the learning curve, and standardise the resulting images seem highly desirable, as they promise to boost data quality and standardisation with minimal human intervention.

The improvement of image acquisition using ML is based on evaluating the current (2D/3D) image on the screen by scoring how closely it resembles the type of view that was intended. This view was learned during a training phase (without explicitly defining the image appearance or content, this is learned by the algorithm). Many ML approaches can be used, but deep learning, using ANN, seems the most promising.

The acquisition of the fetal facial standard plane (FFSP) is a requisite to extract biometric measurements and perform diagnosis during US examination. Lei et al. [8] automated this task with a SVM classifier. More recently, Yu et al. [9] leveraged the power of deep convolutional neural networks (CNNs) to automatically recognize the FFSP during routine US examination. Another standard plane is that of the fetal abdominal region, which allows measuring the abdominal circumference (AC) and estimating fetal weight as a proxy for fetal health. CNNs have already been trained to automatically find the abdominal region in a US image, and then determine image quality by assessing the goodness of depiction for key structures as the stomach bubble and the umbilical vein [10]. In a similar fashion, Rahmutallah et al. [11,12] trained an Adaptative Boosting (AdaBoost) model to detect these two structures in 2D US images for the purpose of scoring image quality. Other ensemble approaches have been proposed to categorise unlabelled fetal 2D US images. In particular, Yaqub et al. used a random forest (RF) classifier to detect meaningful structures from different regions inside the images [13]. A more ambitious project using CNNs targeted the classification of a broader collection of fetal images planes, by automatic recognition of 14 different fetal structures in 2D US images [14]. Concerning 3D fetal US, Raynaud et al. proposed an ensemble of DL for feature extraction and RF for classification of organs with the purpose of automatically encoding anatomical variability while discarding the fetus pose [15].

Detection of the standard scan plane in fetal brain US is an essential step in the assessment of fetal development. This task was achieved by Li et al. [16] using a CNN approach in 3D fetal US. Concerning quality control, Yaqub et al. proposed a DL solution that automatically assessed whether transventricular 2D US images of the fetal brain met clinical standards [17]. Namely, they first localised the fetal brain, detected the regions of interest, and finally learned the US patterns that enable plane verification. ML techniques have also been used to automatically identify the transthalamic plane in 3D US, to then assess brain biometrics such as the fetal biparietal diameter (BPD) and head circumference (HC) [18].

Specific studies involving ML techniques for imaging the fetal heart are still scarce. Among the few examples found, Bridge et al. implemented a framework for tracking the key variables appearing in freehand 2D US scanning videos of the healthy fetal heart, through the use of regression forests [19]. Concerning the electrical activity, Yu et al. used independent component analysis [20] and Muduli et al. used DL to reconstruct the fetal-electrocardiogram (EKG) from abdominal ECG recordings [21]. A next step towards automating the fetal US scanning consists of coupling the image plane/volume recognition with a robot arm that performs the scanning, which has been pioneered by Wang S et al. [22].

ML approaches for improved fetal data acquisition are already a reality in research settings and are expected to become clinically available in the short-term (5 years). In the mid-term, ML techniques may be combined with robotics to automatically extract standardised fetal imaging views.

3. Machine learning for image quantification and feature extraction

Fetal biometric parameters such as HC, BPD, AC, femur length (FL) or thickness of nuchal translucency are commonly used for the estimation of fetal weight, gestational age (GA) and the detection of fetal abnormalities during prenatal US examinations. An accurate estimation of fetal weight and GA is essential to detect any abnormal fetal growth pattern, such as small or large for GA, intrauterine growth restriction (IUGR) or cardiac abnormalities. Kim et al. have recently published a DL model to automatically calculate HC, together with the BPD from 2D US images [23]. A different approach was used by Li et al., which first used RF to localise the fetal head and then ellipse fitting to estimate HC from 2D US images [24]. Van Den Heuvel et al. went a step further and implemented a DL model that calculated HC from obstetric sweep protocol data [25]. These data likely do not contain standard planes, thus their method has a great potential to be applied in resource-constrained countries, where there is a lack of skilled obstetricians. Lorenz et al. have recently published a pipeline combining RF, shape models and CNNs to automatically perform view recognition and anatomical landmark location, with the objective of measuring the AC [26] from 3D US recordings. Similarly, Kim et al. used a CNN to estimate AC from 2D US data [27]. For further information on biometric measurements, we refer the reader to a recent review of automated techniques for the interpretation of fetal abnormalities [28].

ML methods have been proposed in the last decades to improve the estimation of gestational age in women with uncertain or unknown menstrual date [29] and to improve the estimation of fetal weight during gestation. For example, Ashley I et al. [30] explored whether data available at birth can be used to accurately predict estimated fetal weight over the course of gestation using different ML methods such as RF or regression trees in a database of more than 10.000 normal and high-risk pregnancies. The authors found that ML algorithms estimate fetal weight better than other commonly used methods. Chuang et al. [31] developed an ANN to estimate fetal weight using morphometric data from 991 fetuses, reporting a mean absolute percent error of 6.15%.

Apart from measuring fetal biometrics and estimating fetal weight, recent ML approaches have been geared towards segmentation to identify fetal structures and organs to timely find fetal abnormalities so that necessary action can be taken. Namburete et al. used a RF classifier to segment cranial pixels in 2D US images [32]. More recently, Li et al. used a DL approach to automatically segment the fetal body and the amniotic fluid from 2D US data [33]. Other examples of DL for segmentation have targeted the fetal brain and lungs [34,35], and these two organs plus the placenta and the maternal kidneys from magnetic resonance imaging [36]. Last, an ensemble of decision trees has been used to automatically segment fetal brain structures in 3D US images [37].

Concerning the fetal heart, the bulk of research focuses on automatically measuring the heartbeat. Some examples are the detection of cardiac activity from a predefined free-hand US sweep of the maternal abdomen using a classification model [38], the extraction of fetal heart rate from cardiotocograms (CTG) using dimensionality reduction [39], or measuring fetal QRS complexes from maternal ECG recordings using ANN [40]. More recently, Sulas et al. have used ANN to detect heart beats from pulse-wave Doppler envelope signals extracted from B-mode videos [41]. For more on measuring cardiac activity from fetal US using ML techniques, the reader might be interested in the review paper by Alnuaimi et al. [42].

The application of ML algorithms to extract features from fetal echocardiographic data are already being used in some high-end scanners, in particular for the calculation of pulsatility indices from peripheral blood flow recordings. This is expected to be translated to cardiac flows soon. In the mid-term, these scanners will also estimate the GA and assess the fetal growth based on the automatic extraction of the different biometric measurements discussed above.

4. Machine learning for fetal diagnosis

Prenatal diagnosis of fetal abnormalities has greatly benefited from advances in US technology and, in the last years also from the advances in ML. ML algorithms have been used in different applications within fetal US medicine such as to predict preterm births [43,44], risk for euploidy, trisomy 21 and other chromosomal aneuploidies [45] or the prediction of perinatal outcomes on asymptomatic short cervical length [46] among others. Regarding fetal cardiology, one of the subfields in which ML has been extensively applied in the last decades is in the improvement of the diagnosis of fetal hypoxia or acidaemia based on the analysis of CTG. CTG is routinely used to record and monitor fetal heart rate and uterine contractions during antepartum and intrapartum periods, to detect the symptoms of fetal distress as early as possible. In clinical practice, CTG traces are visually examined by clinicians and their interpretation is largely dependent on clinician's expertise leading to high inter and intra-observer variability. Therefore, despite the existence of standardised guidelines, the accuracy and robustness of CTG to improve prenatal outcome remains controversial. The use of ML to improve the predictive capacity of CTG recordings was first presented by Bassil et al. in the late 80's [47]. Since then, several attempts have been made to increase the effectiveness of the automatic evaluation of CTG traced using different ML and DL methods including ANN, SVM, or RF among others. Most of the publications have used two different open access CTG databases to evaluate their proposed ML algorithms: one from the University Hospital in Brno (Czech Republic), including 552 CTG recordings [49]; and another from the University of Porto (Portugal), which includes 2126 CTG recordings [50]. We have summarised the publications on the use of ML in the analysis of CTGs for the last 10 years in the Supplementary Table S1. For a review of older publications, we refer the reader to the review of Graham E.M. et al. [48]. The best results were obtained by Iraji M.S. [51] using the Portuguese database showing an accuracy of 99.5%. There have also been some attempts to translate this into clinical practice by the development of software such as 'Infant', 'PeriCALM' [52,53] or 'Foetos' [54] or the development of mobile/website applications [55,56] to provide additional support in the interpretation of CTG signals and therefore to improve the assessment of fetal status. However, there is no evidence on whether these systems really improved the prediction of fetal distress or acidaemia compared to visual CTG interpretation alone, and reports about their clinical performance were not found. In a recent systematic review, the degree of inter-observer reliability between human and ML interpretation of CTG signals was determined [57], concluding that the use of ML for the interpretation of CTGs during labour does not improve neonatal outcome and has yet to prove its reliability relative to expert observers. The root of the problem may be that any supervised ML-based system needs to be trained with human annotations, and given that the benefit of CTGs themselves for labour monitoring has not been clearly demonstrated, it is not surprising that adding an automatic system to evaluate the CTG signals with similar information does not offer advantages in reducing adverse perinatal outcomes.

IUGR, which affects about 10% of the pregnancies, has been associated with cardiac remodelling in utero that can persists postnatally [58–60]. An early detection of IUGR can improve the perinatal outcome of these fetuses and reduce the risk of cardiovascular mortality in adulthood. The first study proposing the use of ML for the detection of IUGR using biometric data was presented by Gurgen F. et al. in 1997 [61]. In this study an ANN was implemented to approximate the growth curves of fetuses showing an accuracy of 95% in the detection of IUGR. Later, Magenes G. et al.[62] proposed a SVM to detect IUGR using CTG data, showing good classification results in a cohort of 70 fetuses. In 2014, Gadagkar A. et al. [63] developed an ANN system for the diagnosis of IUGR using only 2D US morphometric measurements from almost 300 fetuses, showing similar results that the ones obtained clinically in the same study population. Similarly, Rawat V. et al. [64] implemented an ANN model using again 2D US morphometric measurements from a total of 120 fetuses. Recently, Kuhle S. et al. [65] compared different ML methods to predict fetal growth abnormalities in a cohort of more than 30.000 patients. However, the authors reported that the ML methods used, did not offer any advantage over logistic regression in the prediction of fetal growth abnormalities. The main limitation of all these studies is that the detection of IUGR was performed considering only morphometric data, which only provide information about the fetal weight, without considering any other data such as blood flow velocities or cardiac deformation measured by Doppler or B-mode US respectively. It is known that IUGR fetuses show abnormal blood flow patterns in the fetal circulation detected by Doppler US [66,67], and also signs of longitudinal systolic dysfunction [58]. It has been recently demonstrated that unsupervised ML algorithms using both echocardiographic (including myocardial strain traces) and clinical data can be used to find groups of similar patients within a heart failure cohort and identify individuals with beneficial response to cardiac resynchronization therapy [68]. A similar approach integrating clinical and heterogeneous echocardiographic data could be implemented to improve the detection of

IUGR fetuses, identify those at high risk of adverse perinatal outcome and aid clinicians in finding optimal treatment strategies. However, ML methods require a large number of patients during training in order to be able to capture the range of possible abnormalities, which is a limitation in fetal medicine as the number of patients is scarce. One possibility to overcome this limitation is to combine ML with 'data augmentation' through physiological computational modelling as proposed by Hoodbhoy Z et al [69]. Lumped models of the fetal circulation have demonstrated to be able to realistically simulate the hemodynamics of the fetus in many different conditions [66,67,70], thus providing virtual, but physiologically plausible Doppler traces. Using these models, virtual patients' populations can be created where the ratio of abnormal/normal cases can be increased so that the learning of the ML algorithms is less dependent on the data provided.

Finally, ML has been recently applied to improve the prenatal diagnosis of congenital heart diseases (CHD). Yeo L et al. presented an intelligent navigation method called 'FINE' to automatically obtain different echocardiography anatomical views of the fetal heart and identify abnormalities within the cardiac anatomy. The tool was able to demonstrate evidence of abnormal fetal cardiac anatomy in four abnormal cases [71]. More recently, Arnaout R et al. proposed the use of a fully-convolutional DL method in a supervised manner to: 1) identify the 5 most important views of the fetal heart, 2) segment and measure the cardiac structures, and 3) distinguish between normal and Tetralogy of Fallot and Hypoplastic Left Heart Syndrome (HLHS) using 685 echocardiograms from fetuses from 18 to 24 weeks of GA [72]. The best results were obtained in the diagnosis of HLHS vs. normal with a sensitivity and specificity of 100% and 90%, respectively. Although the results look promising, one of the main limitations of this study is that only two CHDs were evaluated, and that the DL system was only trained with images from one US machine without considering the variability in echocardiographs. Therefore, further studies with bigger datasets from different US machines need to be performed.

5. Conclusions

Given that ML approaches have become ubiquitous in our daily lives, they will become more and more integrated in clinical practice and in the assessment of the fetal heart. It is important to distinguish the different tasks involved in clinical decision making to understand how, and which type of, ML can be optimally employed. For obtaining the best image quality in the shortest possible time and with the smallest learning curve; as well as for the standardised extraction of specific measurements from the images, ML approaches based on Deep Learning have shown great promise and are currently being implemented into the high-end clinical scanners. However, when the diagnostic interpretation is performed, and especially when a treatment decision needs to be made, the 'black-box' approach inherent to e.g. DL becomes problematic given its dependence on a large and very inclusive dataset with correct clinical labels and the inherent difficulty to provide an intuitive clinical explanation for the proposed decision. Here, other ML approaches, based on for example the identification of individuals with similar (complex and multimodal) clinical data and imaging features, seems more promising and is explored in different centres.

Therefore, when carefully used and validated, and taking into account all privacy, security and auditing measures relevant for the use of clinical data, ML can play an important role in standardising fetal cardiac data and provide support in the clinical interpretation and suggestion of the best preventive and interventional approach to optimise perinatal as well as long term cardiovascular health (Figure 3).

6. Statements

6.1. Acknowledgement

No acknowledgements.

6.2. Disclosure Statement

The authors have no conflicts of interest to declare.

6.3. Funding Sources

This project has been partially funded with support of "la Caixa" Foundation under grant agreement LCF/PR/GN14/10270005, the Instituto de Salud Carlos III (PI17/00675) integrados en el Plan Nacional de I+D+I y cofinanciados por el ISCIII-Subdirección General de Evaluación y el Fondo Europeo de Desarrollo Regional (FEDER) "Una manera de hacer Europa", the Centro de Investigación Biomédica en Red de Enfermedades Raras (ERPR04G719/2016), Cerebra Foundation for the Brain Injured Child (Carmarthen, Wales, UK) and AGAUR 2017 SGR grant nº 1531.

7. Author Contributions

Patricia Garcia-Canadilla has participated in the conception and design of the review, drafted the work, approved the final version and agreed on the accuracy and integrity of the work.

Sergio Sanchez-Martinez has participated in the conception and design of the review, drafted the work, approved the final version and agreed on the accuracy and integrity of the work.

Fatima Crispi has participated in the design of the review, revisited it critically for important intellectual content, approved the final version and agreed on the accuracy and integrity of the work.

Bart Bijnens has participated in the conception and design of the review, drafted the work, approved the final version and agreed on the accuracy and integrity of the work.

8. References

1 Rychik J, Tian Z, Bebbington M, Xu F, McCann M, Mann S, et al. The twin-twin transfusion syndrome: spectrum of cardiovascular abnormality and development of a cardiovascular score to assess severity of disease. Am J Obstet Gynecol. 2007 Oct;197(4):392.e1-392.e8.

2 Crispi F, Hernandez-Andrade E, Pelsers MMAL, Plasencia W, Benavides-Serralde JA, Eixarch E, et al. Cardiac dysfunction and cell damage across clinical stages of severity in growth-restricted fetuses. Am J Obstet Gynecol. 2008 Sep;199(3):254.e1-254.e8.

3 Van Mieghem T, Gucciardo L, Doné E, Van Schoubroeck D, Graatsma EM, Visser GHA, et al. Left ventricular cardiac function in fetuses with congenital diaphragmatic hernia and the effect of fetal endoscopic tracheal occlusion. Ultrasound Obstet Gynecol. 2009 Oct;34(4):424–9.

4 Crispi F, Gratacs E. Fetal cardiac function: Technical considerations and potential research and clinical applications. Fetal Diagn Ther. 2012;32(1–2):47–64.

5 Crispi F, Valenzuela-Alcaraz B, Cruz-Lemini M, Gratacós E. Ultrasound assessment of fetal cardiac function. Australas J ultrasound Med. 2013 Nov;16(4):158–67.

6 Mitchell TM. Machine Learning. McGraw Hill; 1997.

7 Deo RC. Machine Learning in Medicine. Circulation. 2015 Nov;132(20):1920–30.

8 Lei B, Zhuo L, Chen S, Li S, Ni D, Wang T. Automatic recognition of fetal standard plane in ultrasound image. 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE; 2014; pp 85–8.

9 Yu Z, Wu L, Ni D, Chen S, Li S, Wang T, et al. Fetal facial standard plane recognition via deep convolutional neural networks. Chinese J Biomed Eng. 2017;36(3):267–75.

10 Wu L, Cheng JZ, Li S, Lei B, Wang T, Ni D. FUIQA: Fetal ultrasound image quality assessment with deep convolutional networks. IEEE Trans Cybern. 2017;47(5):1336–49.

11 Rahmatullah B, Sarris I, Papageorghiou A, Noble JA. Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using AdaBoost. 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2011; pp 6–9.

12 Rahmatullah B, Papageorghiou AT, Noble JA. Image Analysis Using Machine Learning: Anatomical Landmarks Detection in Fetal Ultrasound Images. 2012 IEEE 36th Annual Computer Software and Applications Conference. IEEE; 2012; pp 354–5.

13 Yaqub M, Kelly B, Papageorghiou AT, Noble JA. Guided Random Forests for Identification of Key Fetal Anatomy and Image Categorization in Ultrasound Scans. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham; 2015; pp 687–94.

14 Sridar P, Kumar A, Quinton A, Nanan R, Kim J, Krishnakumar R. Decision Fusion-Based Fetal Ultrasound Image Plane Classification Using Convolutional Neural Networks. Ultrasound Med Biol. 2019 May;45(5):1259–73.

15 Raynaud C, Ciofolo-Veit C, Lefèvre T, Ardon R, Cavallaro A, Salim I, et al. Multi-organ Detection in 3D Fetal Ultrasound with Machine Learning. International Workshop on Fetal and Infant Image Analysis. Springer, Cham; 2017; pp 62–72.

16 Li Y, Khanal B, Hou B, Alansary A, Cerrolaza JJ, Sinclair M, et al. Standard plane detection in 3D fetal ultrasound using an iterative transformation network. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2018;11070 LNCS:392–400.

17 Yaqub M, Kelly B, Papageorghiou AT, Noble JA. A Deep Learning Solution for Automatic Fetal Neurosonographic Diagnostic Plane Verification Using Clinical Standard Constraints. Ultrasound Med Biol. 2017 Dec;43(12):2925–33.

18 Ambroise Grandjean G, Hossu G, Bertholdt C, Noble P, Morel O, Grangé G. Artificial intelligence assistance for fetal head biometry: Assessment of automated measurement software. Diagn Interv Imaging. 2018 Sep DOI: 10.1016/J.DIII.2018.08.001

19 Bridge CP, Ioannou C, Noble JA. Automated annotation and quantitative description of ultrasound videos of the fetal heart. Med Image Anal. 2017 Feb;36:147–61.

20 Yu Q, Yan H, Song L, Guo W, Liu H, Si J, et al. Automatic identifying of maternal ECG source when applying ICA in fetal ECG extraction. Biocybern Biomed Eng. 2018 Jan;38(3):448–55.

21 Muduli PR, Gunukula RR, Mukherjee A. A deep learning approach to fetal-ECG signal reconstruction. 2016 Twenty Second National Conference on Communication (NCC). IEEE; 2016; pp 1–6.

22 Wang S, Housden J, Noh Y, Singh D, Singh A, Skelton E, et al. Robotic-assisted Ultrasound for Fetal Imaging: Evolution from Single-arm to Dual-arm System. 2019 Feb [cited 2019 Aug 21]. Available from: http://arxiv.org/abs/1902.05458

23 Kim HP, Lee SM, Kwon J-Y, Park Y, Kim KC, Seo JK. Automatic evaluation of fetal head biometry from ultrasound images using machine learning. Physiol Meas. 2019 May DOI: 10.1088/1361-6579/ab21ac

24 Li J, Wang Y, Lei B, Cheng J-Z, Qin J, Wang T, et al. Automatic Fetal Head Circumference Measurement in Ultrasound Using Random Forest and Fast Ellipse Fitting. IEEE J Biomed Heal Informatics. 2018 Jan;22(1):215–23.

25 van den Heuvel TLA, Petros H, Santini S, de Korte CL, van Ginneken B. Automated Fetal Head Detection and Circumference Estimation from Free-Hand Ultrasound Sweeps Using Deep Learning in Resource-Limited Countries. Ultrasound Med Biol. 2019 Mar;45(3):773–85.

26 Lorenz C, Brosch T, Ciofolo-Veit C, Tobias K, Lefevre T, Salim I, et al. Automated abdominal plane and circumference estimation in 3D US for fetal screening. In: Angelini ED, Landman BA, editors. Medical Imaging 2018: Image Processing. SPIE; 2018; p 17.

27 Kim B, Kim KC, Park Y, Kwon J-Y, Jang J, Seo JK. Machine-learning-based automatic identification of fetal abdominal circumference from ultrasound images. Physiol Meas. 2018 Oct;39(10):105007.

28 Rawat V, Jain A, Shrimali V. Automated Techniques for the Interpretation of Fetal Abnormalities: A Review. Appl Bionics Biomech. 2018 Jun;2018:1–11.

29 Papageorghiou AT, Kemp B, Stones W, Ohuma EO, Kennedy SH, Purwar M, et al. Ultrasound-based gestational-age estimation in late pregnancy. Ultrasound Obstet Gynecol. 2016 Dec;48(6):719–26.

30 Ashley I. Naimi, Robert W. Platt, Jacob C. Larkin. Machine Learning for Fetal Growth Prediction. Epidemiology. 2018;29(2):290–8.

31 Chuang L, Hwang J-Y, Chang C-H, Yu C-H, Chang F-M. Ultrasound estimation of fetal weight with the use of computerized artificial neural network model. Ultrasound Med Biol. 2002 Aug;28(8):991–6.

32 Namburete AIL, Noble JA. Fetal cranial segmentation in 2D ultrasound images using shape properties of pixel clusters. 2013 IEEE 10th International Symposium on Biomedical Imaging. IEEE; 2013; pp 720–3.

33 Li Y, Xu R, Ohya J, Iwata H. Automatic fetal body and amniotic fluid segmentation from fetal ultrasound images by encoder-decoder network with inner layers. Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS. 2017;1485–8.

34 Rajchl M, Lee MCH, Oktay O, Kamnitsas K, Passerat-Palmbach J, Bai W, et al. DeepCut: Object Segmentation from Bounding Box Annotations Using Convolutional Neural Networks. IEEE Trans Med Imaging. 2017;36(2):674–83.

35 Burgos-Artizzu XP, Perez-Moreno Á, Coronado-Gutierrez D, Gratacos E, Palacio M. Evaluation of an improved tool for non-invasive prediction of neonatal respiratory morbidity based on fully automated fetal lung ultrasound analysis. Sci Rep. 2019 Feb;9(1):1950.

36 Wang G, Li W, Zuluaga MA, Pratt R, Patel PA, Aertsen M, et al. Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning. IEEE Trans Med Imaging. 2018 Jul;37(7):1562–73.

37 Huang R, Xie W, Alison Noble J. VP-Nets : Efficient automatic localization of key brain structures in 3D fetal neurosonography. Med Image Anal. 2018 Jul;47:127–39.

38 Maraci MA, Bridge CP, Napolitano R, Papageorghiou A, Noble JA. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. Med Image Anal. 2017 Apr;37:22–36.

39 Krupa N, MA M, Zahedi E, Ahmed S, Hassan FM. Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. Biomed Eng Online. 2011 Jan;10(1):6.

40 Mantas Lukoševičius ; Vaidotas Marozas. Noninvasive fetal QRS detection using Echo State Network. Computing in Cardiology. 2013; [cited 2019 Jan 24].Available from: https://ieeexplore.ieee.org/abstract/document/6712447

41 Sulas E, Ortu E, Raffo L, Urru M, Tumbarello R, Pani D. Automatic Recognition of Complete Atrioventricular Activity in Fetal Pulsed-Wave Doppler Signals. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2018; pp 917–20.

42 Alnuaimi SA, Jimaa S, Khandoker AH. Fetal Cardiac Doppler Signal Processing Techniques: Challenges and Future Research Directions. Front Bioeng Biotechnol. 2017 Dec;5:82.

43 Rittenhouse KJ, Vwalika B, Keil A, Winston J, Stoner M, Kapasa M, et al. Improving preterm newborn identification in low-resource settings with machine learning. bioRxiv. 2018 May;334904.

44 Hamilton EF, Dyachenko A, Ciampi A, Maurel K, Warrick PA, Garite TJ. Estimating risk of severe neonatal morbidity in preterm births under 32 weeks of gestation. J Matern Neonatal Med. 2018 DOI: 10.1080/14767058.2018.1487395

45 Neocleous AC, Nicolaides KH, Schizas CN. First Trimester Noninvasive Prenatal Diagnosis: A Computational Intelligence Approach. IEEE J Biomed Heal Informatics. 2016 Sep;20(5):1427–38.

46 Bahado-Singh RO, Sonek J, McKenna D, Cool D, Aydas B, Turkoglu O, et al. Artificial Intelligence and amniotic fluid multiomics analysis: The prediction of perinatal outcome in asymptomatic short cervix. Ultrasound Obstet Gynecol. 2018 Oct DOI: 10.1002/uog.20168

47 Bassil HE, Dripps JH. Real time processing and analysis of fetal phonocardiographic signals. Clin Phys Physiol Meas. 1989 [cited 2018 Nov 13]. ;10 Suppl B:67–74.

48 Graham EM, Petersen SM, Christo DK, Fox HE. Intrapartum Electronic Fetal Heart Rate Monitoring and the Prevention of Perinatal Brain Injury. Obstet Gynecol. 2006 Sep;108(3, Part 1):656–66.

49 Chudáček V, Spilka J, Burša M, Janků P, Hruban L, Huptych M, et al. Open access intrapartum CTG database. 2014; [cited 2019 Jun 13].Available from: http://www.biomedcentral.com/1471-2393/14/16DATABASE

50 Ayres-de-Campos D, Sousa P, Costa A, Bernardes J. Omniview-SisPorto® 3.5 – a central fetal monitoring station with online alerts based on computerized cardiotocogram+ST event analysis. J Perinat Med. 2008 Jan;36(3):260–4.

51 Iraji MS. Prediction of fetal state from the cardiotocogram recordings using neural network models. Artif Intell Med. 2019 May;96:33–44.

52 Brocklehurst P. A study of an intelligent system to support decision making in the management of labour using the cardiotocograph – the INFANT study protocol. BMC Pregnancy Childbirth. 2016 Dec;16(1):10.

53 Hamilton EF, Warrick PA. New perspectives in electronic fetal surveillance. J Perinat Med. 2013 Jan [cited 2019 Aug 21]. ;41(1):83–92.

54 Alonso-Betanzos A, Moret-Bonillo V, Hernandez-Sande C. Foetos: an expert system for fetal assessment. IEEE Trans Biomed Eng. 1991;38(2):199–211.

55 Akbulut A, Ertugrul E, Topcu V. Fetal health status prediction based on maternal clinical history using machine learning techniques. Comput Methods Programs Biomed. 2018 Sep;163:87–100.

56 Kazantsev A, Ponomareva J, Kazantsev P, Digilov R, Ping Huang. Development of e-health network for in-home pregnancy surveillance based on artificial intelligence. Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics. IEEE; 2012; pp 82–4.

57 Balayla J, Shrem G. Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis. Arch Gynecol Obstet. 2019 May;1–8.

58 Crispi F, Bijnens B, Figueras F, Bartrons J, Eixarch E, Le Noble F, et al. Fetal growth restriction results in remodeled and less efficient hearts in children. Circulation. 2010 Jun;121(22):2427–36.

59 Cruz-Lemini M, Crispi F, Valenzuela-Alcaraz B, Figueras F, Sitges M, Bijnens B, et al. Fetal cardiovascular remodeling persists at 6 months in infants with intrauterine growth restriction. Ultrasound Obstet Gynecol. 2016 Sep;48(3):349–56.

60 Sarvari SI, Rodriguez-Lopez M, Nuñez-Garcia M, Sitges M, Sepulveda-Martinez A, Camara O, et al. Persistence of Cardiac Remodeling in Preadolescents With Fetal Growth Restriction. Circ Cardiovasc Imaging. 2017 Jan;10(1). DOI: 10.1161/CIRCIMAGING.116.005270

61 Gurgen F, Onal E, Varol FG. IUGR detection by ultrasonographic examinations using neural networks. IEEE Eng Med Biol Mag. 1997;16(3):55–8.

62 Magenes G, Pedrinazzi L, Signorini MG. Identification of fetal sufferance antepartum through a multiparametric analysis and a support vector machine. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2004; pp 462–5.

63 Gadagkar A V., Shreedhara KS. Features Based IUGR Diagnosis Using Variational Level Set Method and Classification Using Artificial Neural Networks. 2014 Fifth International Conference on Signal and Image Processing. IEEE; 2014; pp 303–9.

64 Rawat V, Jain A, Shrimali V, Rawat A. Automatic Detection of Fetal Abnormality Using Head and Abdominal Circumference. Lecture Notes in Computer Science (LNCS, volume 9876). Springer, Cham; 2016; pp 525–34.

65 Kuhle S, Maguire B, Zhang H, Hamilton D, Allen AC, Joseph KS, et al. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. BMC Pregnancy Childbirth. 2018 Dec;18(1):333.

66 Garcia-Canadilla P, Rudenick PA, Crispi F, Cruz-Lemini M, Palau G, Camara O, et al. A Computational Model of the Fetal Circulation to Quantify Blood Redistribution in Intrauterine Growth Restriction. PLoS Comput Biol. 2014 Jun;10(6):e1003667.

67 Garcia-Canadilla P, Crispi F, Cruz-Lemini M, Triunfo S, Nadal A, Valenzuela-Alcaraz B, et al. Patient-specific estimates of vascular and placental properties in growth-restricted fetuses based on a model of the fetal circulation. Placenta. 2015 Jul;36(9):981–9.

68 Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, et al. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. Eur J Heart Fail. 2019 Jan;21(1):74–85.

69 Hoodbhoy Z, Hasan B, Jehan F, Bijnens B, Chowdhury D. Machine learning from fetal flow waveforms to predict adverse perinatal outcomes: a study protocol. Gates open Res. 2018 Feb;2:8.

70 Mínguez PG, Bijnens B, Bernardino G, Lluch È, Soveral I, Gómez O, et al. Assessment of haemodynamic remodeling in fetal aortic coarctation using a lumped model of the circulation. 2017. DOI: 10.1007/978-3-319-59448-4_45

71 Yeo L, Romero R. Fetal Intelligent Navigation Echocardiography (FINE): a novel method for rapid, simple, and automatic examination of the fetal heart. Ultrasound Obstet Gynecol. 2013 Sep;42(3):268–84.

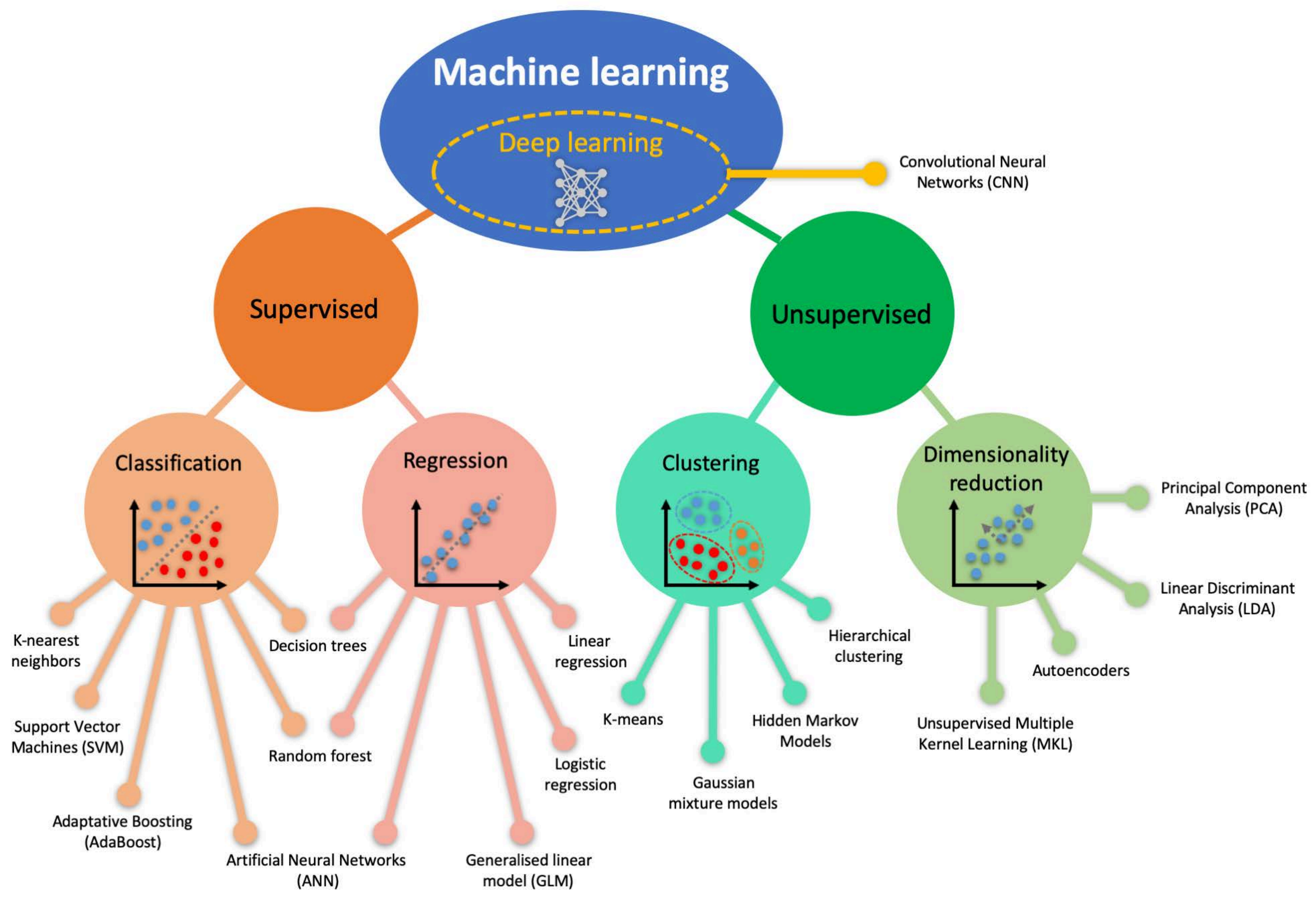
72 Arnaout R, Curran Mbbs Bsc L, Chinn E, Zhao Y, Rdcs P, Moon-Grady A. Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions. 2018. DOI: arXiv:1809.06993v1

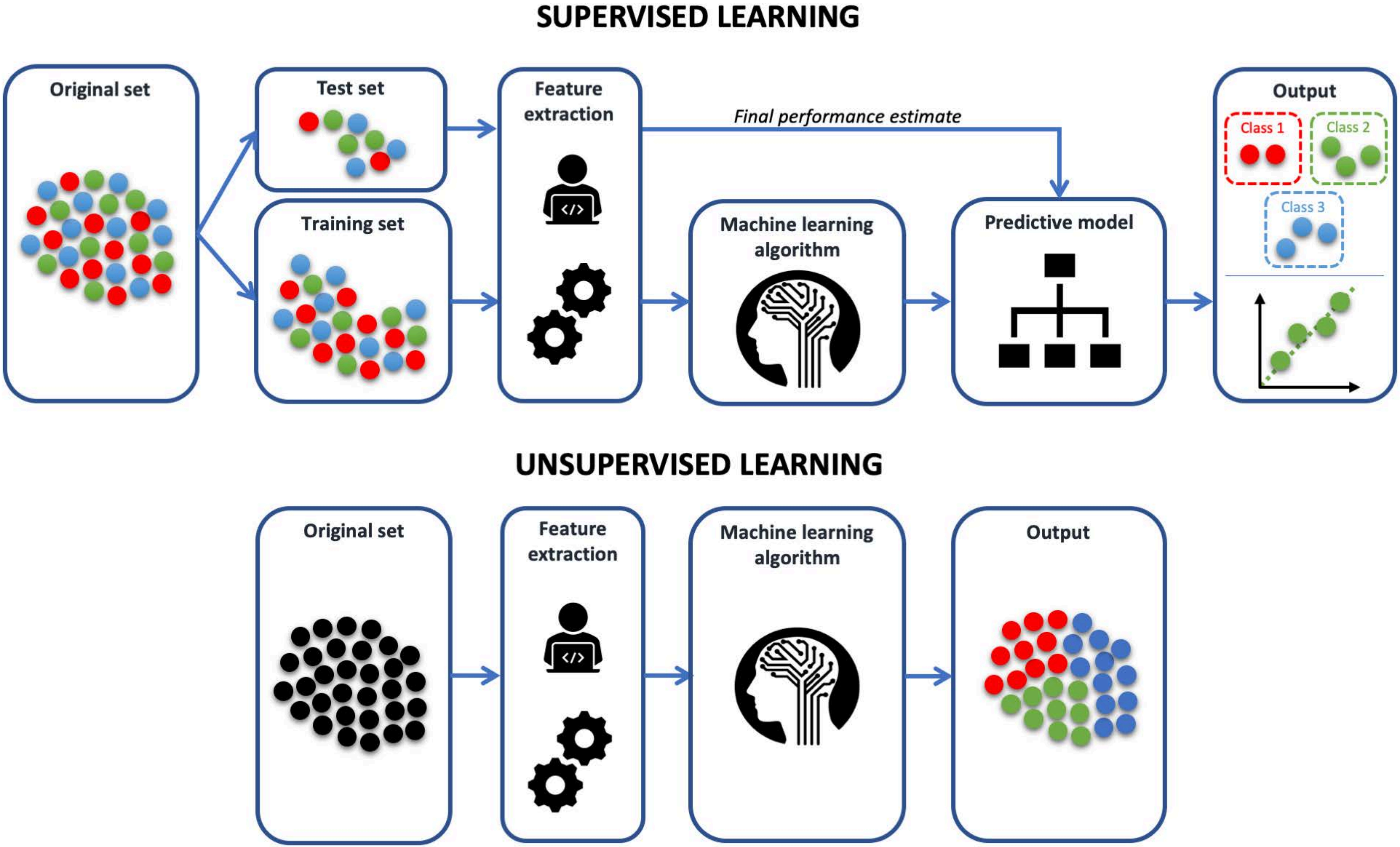
9. Figure Legends

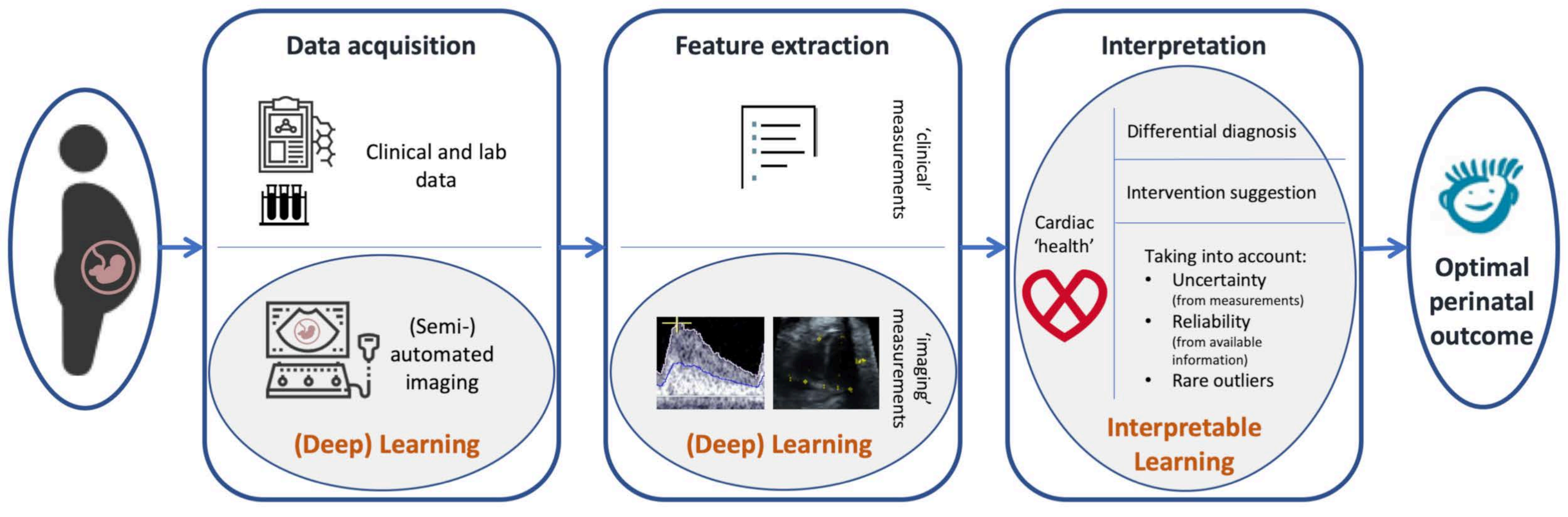
Figure 1. Classification of machine/deep learning algorithms. Deep learning is a subset of machine learning based on artificial neural networks and can be applied in a supervised or unsupervised manner.

Figure 2. Pipeline of supervised (top) and unsupervised (bottom) learning applications.

Figure 3. Overview of the application of machine/deep learning in fetal cardiology. In the short term, ML will help the sonographer to acquire a full set of optimal and standardised images in the shortest possible time. This can improve data quality and interpretability. Next, ML will be used for the (semi-) automated extraction of features (measurements) from the images, thus again improving standardisation and efficiency of the imaging department. These first component, essential for the use of images in fetal cardiology, will likely benefit greatly from the development in deep learning. The next step in clinical decision making is the data interpretation for diagnosis and therapy planning. This component is much riskier so that interpretability and reliability of the ML decision support becomes a crucial factor. Therefore, in the foreseeable future, this will stay fully in the hands of the clinician but ML can provide helpful support by presenting the data in such a way that the comparison of an individual patient with knowledge from patho-physiology and clinical trials/research becomes an easier task when a huge amount of complex data (from anamnesis to images over lab results..) is available.


Machine learning
Deep learning
Convolutional Neural Networks (CNN)
Supervised
Unsupervised
Classification
Regression
Clustering
Dimensionality reduction
K-nearest neighbors
Decision trees
Support Vector Machines (SVM)
Random forest
Adaptative Boosting (AdaBoost)
Artificial Neural Networks (ANN)
Linear regression
Logistic regression
Generalised linear model (GLM)
K-means
Gaussian mixture models
Hidden Markov Models
Hierarchical clustering
Principal Component Analysis (PCA)
Linear Discriminant Analysis (LDA)
Autoencoders
Unsupervised Multiple Kernel Learning (MKL)


SUPERVISED LEARNING
Original set
Test set
Training set
Feature extraction
Final performance estimate
Machine learning algorithm
Predictive model
Output
Class 1
Class 2
Class 3
UNSUPERVISED LEARNING
Original set
Feature extraction
Machine learning algorithm
Output



Data acquisition
Clinical and lab data
(Semi-) automated imaging
(Deep) Learning
Feature extraction
'clinical' measurements
'imaging' measurements
(Deep) Learning
Interpretation
Cardiac 'health'
Differential diagnosis
Intervention suggestion
Taking into account:
• Uncertainty (from measurements)
• Reliability (from available information)
• Rare outliers
Interpretable Learning
Optimal perinatal outcome