UNIVERSITA' POLITECNICA DELLE MARCHE
FACOLTA' DI INGEGNERIA

Corso di Laurea magistrale in
BIOMEDICAL ENGENEERING

Tesi di laurea:

UNSUPERVISED CLASSIFICATION OF HYPOXIC FETUSES

Relatore:
Prof.ssa **Burattini Laura**

Correlatore:
Dott.ssa **Sbrollini Agnese**

Tesi di Laurea di:
Ravallese Anna Gaia

A.A. **2021** / **2022**

ABSTRACT

Fetal hypoxia is characterized by an inadequate delivery of oxygen to the baby's tissues in the womb. Long-term oxygen deprivation may cause damage to the central nervous system, culminating in cerebral palsy and, in the worst instances, death. The fetus exchanges an exact concentration of oxygen and carbon dioxide with the mother through the placenta. If this does not take place, the fetus might develop fetal hypoxia. Cardiotocography (CTG) is the most extensively used screening tool for fetal hypoxia. CTG is a test that combines cardiography and tocography to detect the fetal heart rate (FHR) and uterine contractions (UC) at the same time. The purpose of this study is to assess CTG tracings using an unsupervised deep neural network in order to eliminate human error when determining whether or not certain tracings are hypoxic.

To do this, many studies that use unsupervised learning to analyze CTG tracings in order to identify fetal hypoxia were compared. Research was conducted on the most current studies, including those that compared deep learning techniques to machine learning methods. All research evaluated demonstrate that deep learning techniques get the greatest outcomes. One of the primary reasons unsupervised deep learning was chosen was because it most closely resembles human behaviour, making it ideal for replacing obstetricians and preventing human eye mistake.

This kind of research utilizes a database obtained from the Faculty of Medicine at the University of Porto, which has 21 characteristics used to evaluate FHR and UC. It covers the tracings of 2126 third-trimester pregnant women. Expert obstetricians assessed CTGs using three categories: normal, suspicious, and pathological. This categorization served as a reference for comparing the results generated from the neural network used in this study.

The database was split into two datasets, one for training (70%) and one for testing (30%), while maintaining a uniform distribution of classes. Neural Net Clustering was utilized as a Matlab toolbox to produce a Self-Organizing Map (SOM) for both training and testing in order to develop the neural network. SOMs are a type of neural network organization trained using unsupervised learning to produce a representation of training samples in a two-dimensional space where output neurons are organized in 2D grids and each input is connected to all output neurons preserving the topological properties of

the input space. Then, a quantization vector was employed to determine the class partitioning. Finally, the predicted classes were compared with the real classes.

The network already in training phase achieves an accuracy of 74%, which increases throughout the test phase to reach 82%, so the results that have been produced so far are quite encouraging. These findings are quite promising since they clear the way for more advancements and improvements in the future. In fact, the objective for the near future is to make this method so self-sufficient and effective that it can be used in hospitals. This will allow medical staff to be trained on the method, which will in turn reduce the risks that a situation of fetal hypoxia may cause.

CONTENTS

CHAPTER ONE: FETAL DISTRESS CLASSIFICATION

1.1 Introduction

Hypoxia and perinatal asphyxia are responsible for nearly one-third of all neonatal deaths. Thus, it is a frequent condition that should not be underestimated. This is the reason why there are numerous studies on fetal hypoxia. This chapter will describe the physiology of the fetus and will focus on the fetal development of the hypoxic condition and all that this situation implies: the causes, the pathophysiology of this process and the defence mechanisms adopted by the fetus to counteract this phenomenon will be investigated.

Fetal hypoxia indicates an insufficient supply of oxygen to the tissues of the baby in the womb. If the baby does not receive enough oxygen for a long period of time, damage to the central nervous system can occur, leading to cerebral palsy or, in the worst cases, death.

Fortunately, the fetus has a compensatory mechanism that allows it to defend itself from the condition of hypoxia. This safety margin, however, is rather limited and inherent to the time in which the hypoxic condition occurs: fetal hypoxia cannot last long without causing permanent damage to the fetus. Depending on the severity and duration of the condition of oxygen insufficiency or absence, different degrees of hypoxia can be distinguished: from acute to chronic.

Many studies state that fetal cardiac activity is the main source of information about fetal health and in particular the detection of fetal hypoxia. Several intra-partum surveillance methods have been employed to detect signs of hypoxia more readily so that the risk of cerebral palsy can be minimized and the mortality rate among new-borns can be reduced. Two most widely used intrapartum monitoring techniques are intermittent auscultation (IA) and the Doppler machine [1]. But since 1960, cardiotocography (CTG) has replaced all other traditional methods of monitoring the Fetal Heart Rate (FHR).

1.2 Fetal distress

Through the placenta, the fetus exchanges oxygen and carbon dioxide with the mother. The exchange is based on precise blood gas concentration, uterine blood supply, and fetal gas transport. A lack of any of these factors can cause fetal hypoxia, which can lead to acidosis [2].

"Fetal distress" is a term used to describe a fetus in which both oxygen deprivation and carbon dioxide accumulation are evident. This situation can degenerate into a condition of hypoxia and acidosis in the period before labor or during birth.

Fetal distress is a widely used although poorly defined term in the literature. It refers to a condition of hypoxia that can lead to irreparable brain damage if not interrupted as soon as possible.

In 2005, the American College of Obstetricians and Gynecologists (ACOG) recommended changes to the terms "fetal distress" and "birth asphyxia" because they were considered imprecise and nonspecific. In fact, it is preferred to use the term "non-reassuring fetal status", followed by a more in-depth description, to refer to suspicious fetal cardiac tracings. This is because most "non-reassuring tracings" end in the birth of healthy infants. Birth asphyxia, on the other hand, should be described with reference to its specific time to allow its severity to be assessed.

Below are some basic definitions inherent to the topic:

- Hypoxemia: decrease of oxygen in blood
- Hypoxia: decrease in tissue oxygen level
- Acidemia: increase in the concentration of hydrogen ions in the blood
- Acidosis: increased concentration of hydrogen ions in the tissue
- Asphyxia: general lack of oxygen affecting priority organs

It is not easy to identify a priori the severity of damage caused by fetal hypoxia, especially if the duration of this condition is unknown.

The most serious consequences may be:

- cerebral paralysis and mental disabilities
- fetal death

Fetal hypoxia can be classified as acute hypoxia or chronic hypoxia depending on the stage of intra-uterine fetal life. The former usually occurs during the process of labor, whereas the latter can occur throughout the pregnancy [2].

In Fig. 1 it is observable a scheme that helps to understand in which occasions may occur fetal distress:

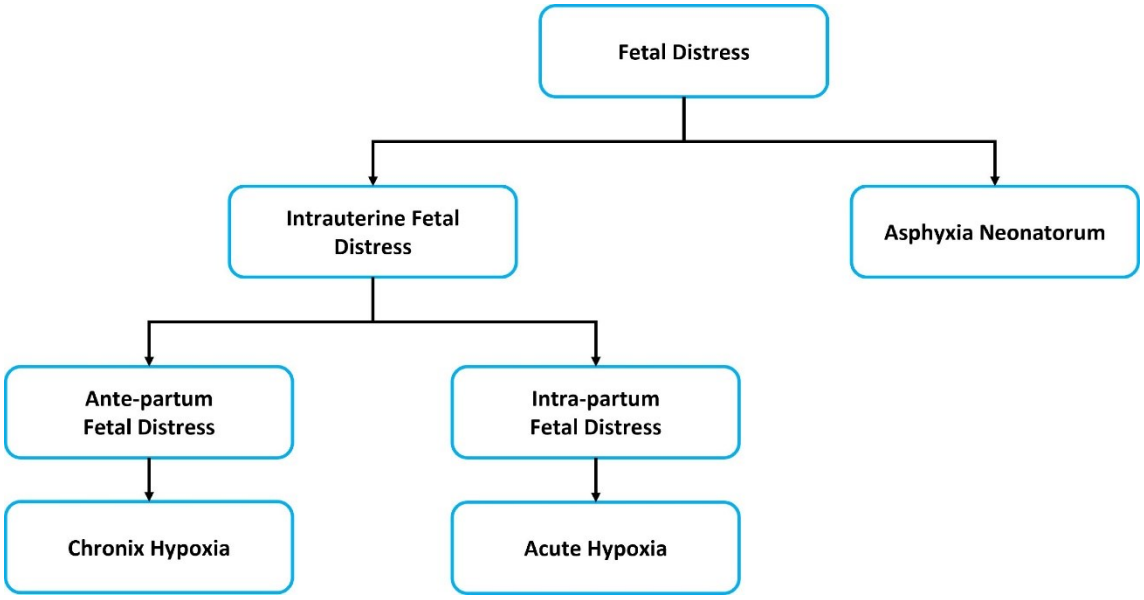

Fig. 1: scheme of fetal distress.

1.3 Fetal heart circulation

In order to have a clearer view of the situations that can lead the fetus to a hypoxic condition, it is preferable to know what is the physiology of the fetus pre- and post-natal.

Fig. 2 shows the blood circulation of the human being respectively before and after birth, which are described below.

- *Prenatal development:*

The transition from embryo to fetus in humans occurs between the 9th and 10th weeks of gestation. Pregnancy can last from 38 to 40 weeks, and during this period the fetus must be continuously monitored to prevent it from developing malformations.

During intrauterine life, the lungs of the fetus are not yet in use, so the gaseous exchanges, the intake of nutritious materials and the elimination of catabolites are carried out through the placenta, a fetal annex created during the 4th week of pregnancy and expelled after delivery.

The placenta provides the fetal blood with a continuous supply of oxygen and the ability to eliminate carbon dioxide and metabolic materials.

The heart and blood vessels of the circulatory system are already formed during embryonic development but continue to develop during the fetal phase.

Blood is transported from the placenta to the fetus by the umbilical vein. Part of the blood enters the ductus venosus and flows into the inferior vena cava, while the other part reaches the liver.

The umbilical vein that supplies the liver, first joins the right portal vein and then reaches the inferior vena cava through the right hepatic vein. From here, the blood reaches the right atrium of the heart. In the fetus, there is an opening between the right and left atrium known as the *foramen ovale*. Because of this, most of the blood flows from the right to the left atrium, bypassing the pulmonary circulation. Most of the blood flow arrives in the left ventricle where it is pumped to the rest of the body through the aorta. At this point, some of the blood from the aorta reaches the umbilical arteries through the iliac arteries and re-enters the placenta. Here, waste products from the fetus, including carbon dioxide, are absorbed and enter the mother's circulation.

As can be seen on Fig. 2, some of the blood from the right atrium does not enter the left atrium but enters the right ventricle to be pumped into the pulmonary artery. This is because in the fetus there is the ductus arteriosus, which connects the pulmonary artery to the aorta, which is intended to direct this blood away from the lungs, which are not used for respiration as the fetus is suspended in amniotic fluid [3].

- *Postnatal development:*

After birth, the system changes with the baby's first breath. Blood from the lungs is directed through the pulmonary veins to the left atrium, producing an increase in pressure that causes the foramen ovale to close, completing the separation of the circulatory system. When the foramen ovale closes, it will be called the *fossa ovalis*. The ductus arteriosus closes within two days after birth, leaving the arterial ligament; the umbilical vein and the venous duct close by the 5th day after birth, leaving the ligamentous round and the venous ligament, respectively.

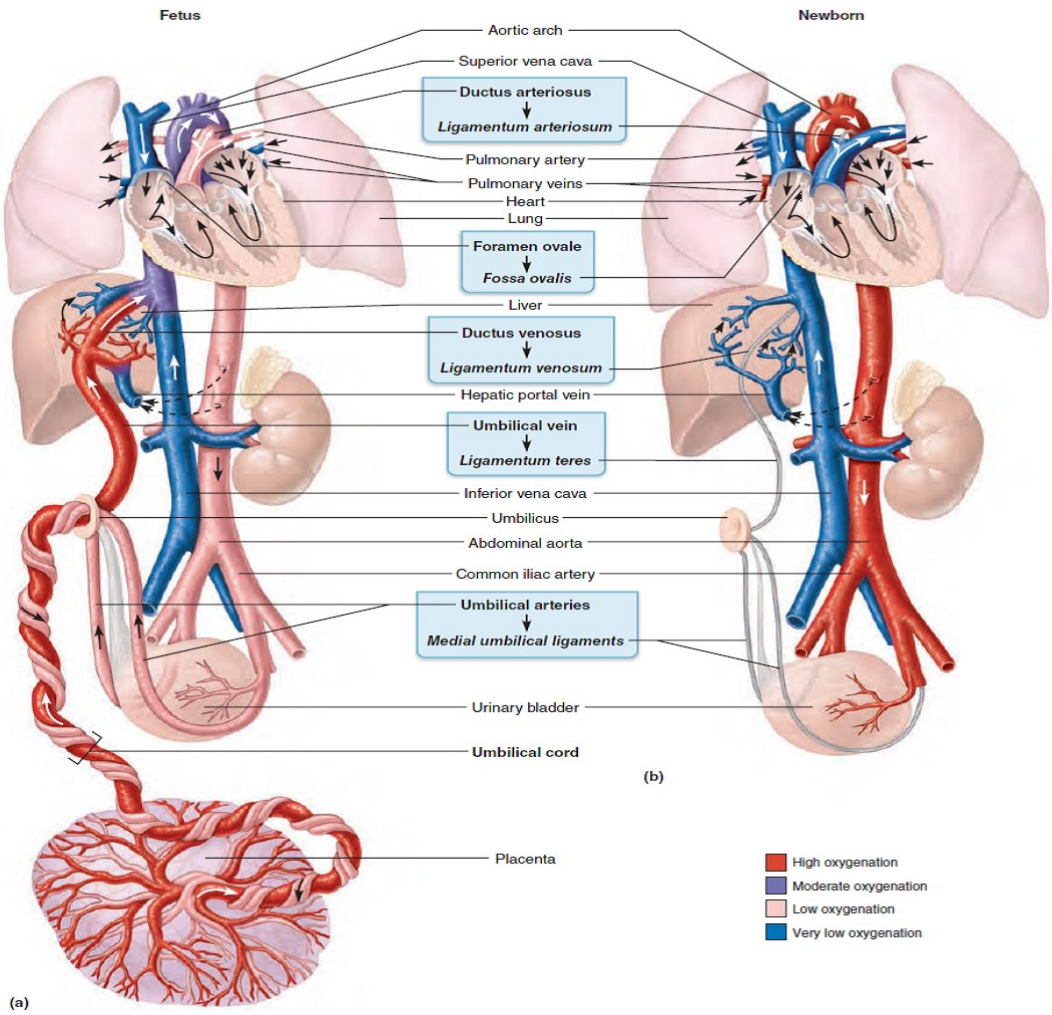

Fig. 2: fetal (a) and neonatal (b) circulation.

1.4 Fetal distress pathophysiology

It is possible to refer to fetal hypoxia in three situations:

- maternal blood oxygenation is impaired

- the perfusion of the placenta is reduced

- the transmission of sufficiently oxygenated blood from the placenta to the fetus does not occur: the blood arrives but is not sufficiently oxygenated. In Fig. 3 an illustration of the blood flow from placenta to the brain in condition of hypoxia.

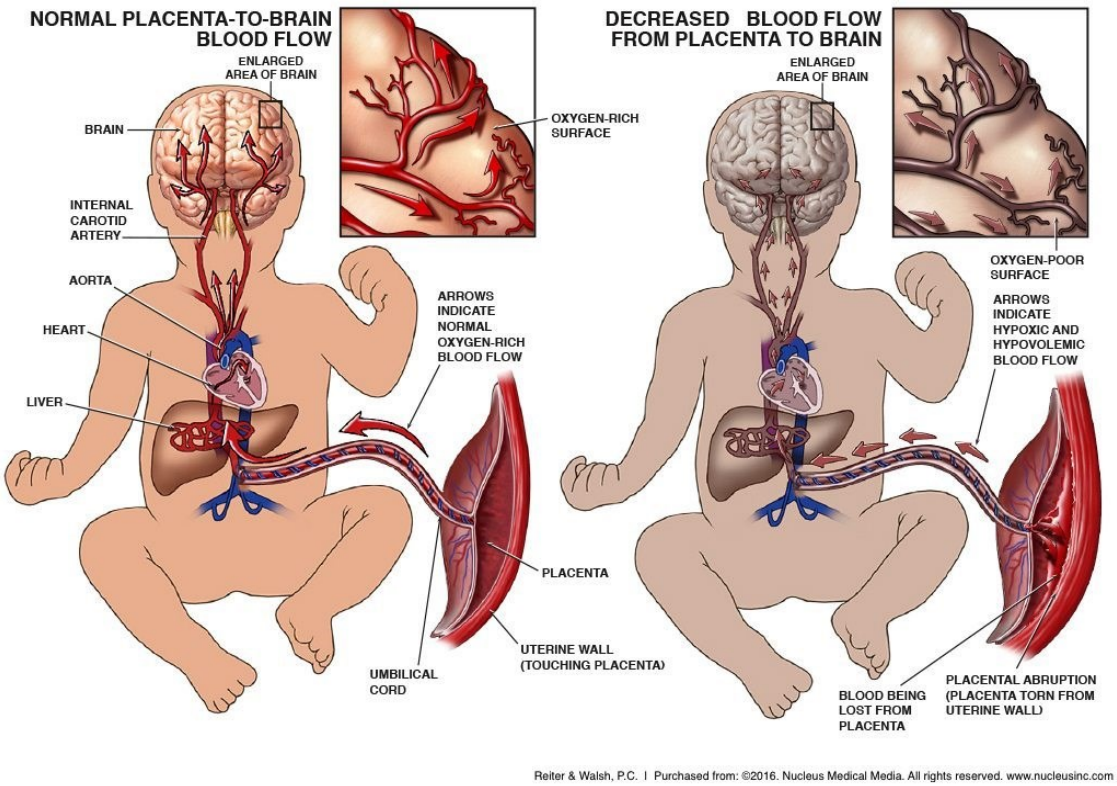

Fig. 3: differences from normal and decreased blood flow from placenta to brain.

In all cases in which adequate oxygenation of the fetus does not occur, the so-called "anaerobic metabolisms" occur, characterized by the production of organic acids including lactic acid.

Lactic acid is very important because in high amounts it can cause metabolic acidosis with low fetal pH and therefore fetal distress [4].

However, there are compensatory mechanisms that the fetus applies to adapt to the condition of oxygenation deficiency in which it finds itself.

The first compensatory mechanism is the one that helps the fetus to release more oxygen to the tissues: considering that the concentration of oxygen in the umbilical vein (40 mmHg) is lower than that of the maternal artery (95 mmHg), the condition occurs in which there is a higher concentration of hemoglobin in the fetal blood (in fact, the oxygen saturation in the umbilical vein is very similar to that of maternal arterial blood). This is due to one of the first compensatory mechanisms that the fetus puts in place to protect its own well-being.

Another fundamental compensatory mechanism applied by the fetus to cope with low blood oxygen concentration is the redistribution of blood to priority organs and increased oxygen extraction from tissues [4]. In fact, the fetus is able to limit hypoxic effects by increasing the supply of nutrient-rich and oxygen-rich blood to the most vital organs such as the heart, brain, and upper body and by reducing the perfusion of the lower extremities, kidneys, and gastrointestinal tract [5].

If the redistribution of oxygen in the blood occurs for a long time in the antepartum phase, there may be not insignificant consequences for the fetal heart. Indeed, a substantial reduction in fetal oxygenation leads to a reduction in cardiac output and predisposes the child to the development of arterial hypertension. There is a study in this regard conducted by Barker et al [6] that found that infants who suffered from fetal hypoxia are likely to develop severe cardiovascular disease in the future. This happens because the physiological readjustment of the fetus, put in place for its survival, implies a delay in the development of key organs that will lead to pathological conditions after birth.

The pathophysiology of fetal hypoxic disorder can be observed in Fig. 4.

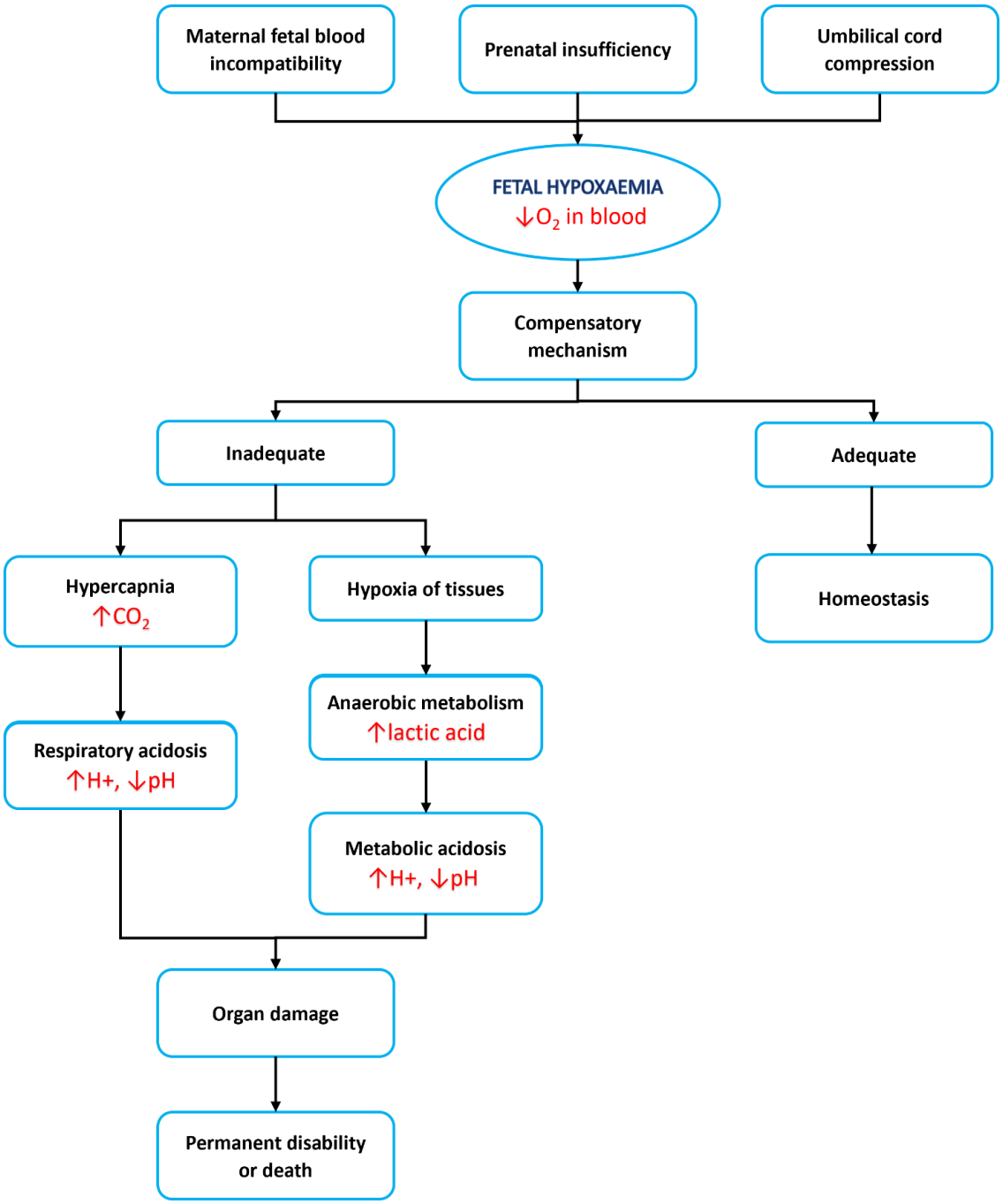

Fig. 4: fetal distress pathophysiology.

As shown in Fig. 4, the response to fetal hypoxia is based on a compensatory mechanism that may be adequate or inadequate. In the former case, the fetus has an adequate blood oxygenation and achieves homeostasis. In the second case, however, the fetus may suffer from dysfunction such as an increased level of carbon dioxide in the blood, which can lead to respiratory acidosis. The increase in carbon dioxide in the blood is due to a reduction in oxygen perfusion within the blood, which then causes an increase in the partial pressure of CO_2 and a reduction in pH, which then leads to respiratory

acidosis [6]. Another condition that can occur if the compensatory mechanism is inadequate is the hypoxia of the tissue, as a result of which there will be the activation of anaerobic metabolism (with consequent formation of lactic acid) and metabolic acidosis. Both situations eventually lead to damage to the organs of the fetus with the risk that the fetus may develop disabilities or may die. In fact, severe respiratory or metabolic acidosis can lead to death because it implies damage to several organs including mainly heart, lungs and kidneys. Therefore, in order to correctly diagnose fetal asphyxia at birth there must be, in addition to metabolic acidosis, evidence of malfunction of one of these organs [7].

1.5 Fetal hypoxia causes

Fetal hypoxia, whether chronic (whether it lasts for days or weeks) or acute (whether it lasts for minutes or hours), has in both cases various factors that affect and reduce normal blood oxygen levels. As shown in Table 1, the causes of fetal hypoxia that lead to acidosis can be classified as follows: maternal, placental, or fetal causes. The consequences worsen with duration [8].

ACUTE FETAL HYPOXIA CAUSES:		
Maternal	**Placental**	**Fetal**
• Hypotension or hypovolemia • Hemorrhage • Vaso-vagal attack • Epidural anesthesia • Uterine contractions	• Placental abruption • Hyperstimulation secondary to oxytocin, prostaglandins E2	• Umbilical cord compression • Oligohydramnios: low amniotic fluid levels • Umbilical cord prolapse
CHRONIC FETAL HYPOXIA CAUSES:		
Maternal	**Placental**	**Fetal**
• Severe respiratory or cardiac disease • Connective tissue diseases • Significant anemia • Antiphospholipid syndrome	• Inadequate trophoblastic invasion of the myometrium during the early stages of pregnancy • Utero-placental dysfunction • Placenta previa	• Rh factor disease anemia • Parvovirus infection • α-thalassemia • Feto-maternal hemorrhage • Severe structural abnormalities of the heart

Table 1: Causes of fetal hypoxia.

In the first case of acute hypoxia with maternal causes, one has that phenomenon such as hypotension and hemorrhage in each case cause a reduction of maternal blood supply; in the case of uterine contractions, on the other hand, they can also interrupt uterine blood flow by causing an increase in fetal pressure. In case of placental causes, the detachment can disrupt utero-placental circulation by tearing uterine arteries from the placenta. Regarding fetal causes, blood flow from the placenta to the fetus is often

compromised by compression of the umbilical cord during delivery, and this can sometimes occur even before labor if there is a knot in the cord. In addition, because the fetus can compensate by increasing oxygen extraction, blood flow to the fetus must be reduced by at least 50% to cause hypoxia [9].
Low amniotic fluid levels (oligohydramnios) can also lead to hypoxia.

In case of chronic hypoxia due to maternal causes, severe respiratory or cardiac disease reduces maternal blood oxygenation levels; connective tissue diseases such as systemic lupus erythematosus and pre-eclampsia reduce blood flow to the placenta. Also, the inadequate trophoblastic invasion of the myometrium can reduce maternal blood supply and so oxygen delivery to the uterus.

1.6 Diagnosis of fetal hypoxia

Since early diagnosis is important, as in many cases of disease, for the health of the new-born, various methods of intra-partum fetal surveillance are employed in order to minimize the risk of cerebral palsy and reduce the mortality rate. As pointed out in the previous paragraphs, the compensation mechanism is the first form of defence that the fetus puts in place to preserve the oxygenation of vital organs such as the brain and heart; therefore, numerous changes in the cardio-respiratory system occur. Considering, thus, that the fetal cardiac system can be modified as a response to hypoxia, the presence of this phenomenon may itself constitute a sign of a state of fetal distress and hence hypoxia [10].

Various methods have been tested over the years considering also different parameters involved in creating a fetal biophysical profile. One such method was proposed by Manning et al [11].

Manning assessed general fetal activity by evaluating five biophysical variables: fetal respiratory movements, fetal movements, fetal tone, qualitative amniotic fluid volume and the non-stress test to form a homeostatic profile of the fetus. However, this approach while theoretically complete, provided a low number of false negatives and a high number of false positives (as high as >50%). One study to consider is that of Thacker et al. [12], who showed that fetal conditions are more accurately diagnosed when considering a set of measurements of all biophysical variables as opposed to measuring each variable individually.

One test that has been performed over the years but is quite invasive is cordocentesis: a needle collection of fetal blood from the umbilical cord under ultrasound guidance, usually used to detect hypoxia due to abnormalities in placental development [13].

It is also of particular importance to recognize a condition of fetal hypoxia especially during labor/delivery, as fetuses are subjected to maximum stress, which puts their health at risk. About three-quarters of infants with severe hypoxic-ischemic encephalopathy (HIE) die from multiple organ failure. Not all of them survive, and those who do, have consequences such as cerebral palsy or severe mental retardation [14]. Continuous monitoring of labor is therefore essential to ensure fetal well-being.

Many studies indicate that fetal cardiac activity is the main source of information on fetal health and specially to detect fetal hypoxia [15]. Indeed, among today's prenatal screening methods are primarily FHR monitoring using CTG and IA.

CHAPTER TWO: CARDIOTOCOGRAPHY

2.1 Introduction

Alan Bradfield, Orvan Hess, and Edward Hon are the three clinicians credited with inventing fetal monitoring. Later on, in the early 1960s, Hewlett Packard and Konrad Hammacher proposed the concept for the cardiotocograph, which is a more advanced kind of non-invasive fetal monitoring that takes place during the antepartum period.

Cardiotocography is a non-invasive screening procedure that can be done beginning in the 36th week of pregnancy. Its purpose is to identify fetuses that are affected by more or less severe hypoxia, which can ultimately result in death.

Cardiotocography is an examination that utilizes *cardiography* and *tocography*, two subspecialties of medicine and obstetrics, in order to monitor the fetal heartbeat and uterine contractions at the same time.

The technique of cardiography makes it possible to characterize the elevations and depressions of the chest wall in connection to the rhythmic motions of the heart.

Tocography makes it feasible to study uterine contractions since it provides information on the strength of these contractions. As a result, tocography makes it possible to discover any and all abnormalities of the uterine muscle's ability to contract.

In instance, a state of excessive contractility known as hyperkinesis, in which uterine activity becomes excessively intense, is known as hypokinesis. Hypokinesis is associated with quite extended labor periods and may suggest that the fetus is experiencing distress.

Monitoring of FHR during labor is now considered to be standard practice in the medical world.

The cardiotocographic examination gives information about:

- the regular beats per minute [bpm] of FHR (usually between 110 and 160 bpm)
- accelerations (transient increases in heart rate)
- decelerations (transient decreases in heart rate).

This is very significant because the fetal heart rate can be altered by certain aspects of the labor and delivery process, such as the contractions that the mother goes through or the movements that she makes.

The sensitivity of the fetal CTG test is quite high, but the specificity of the test is not particularly high. In point of fact, it has a false-negative rate of 5%, which indicates that fetal distress is rarely missed; consequently, a normal tracing is a good guarantee of fetal well-being when there are no abnormalities found in it. However, it has a false-positive rate of forty percent, and as a consequence, it may present a picture that is unnecessarily alarming even when the fetus is doing exceptionally well.

However, it has also contributed to an increase in the number of cesarean deliveries performed for suspected fetal distress. The introduction of cardiotocography has led to a significant decrease in the number of cases of fetal distress and deaths that occur during labor. When used in pregnancies that are considered to be “at risk,” CTG is extremely useful for diagnostic purposes, therapeutic purposes, and even medico-legal purposes.

On the right side of the Fig. 5 there is an illustration of a cardiotocograph, and on the left side is a depiction of how it interfaces to the womb in order to determine the baby’s FHR as well as the mother’s uterine contractions (UC). In order to accomplish its objective, it is fitted with two distinct kinds of sensors, each of which is attached to the womb. On the right side it is also possible to see the monitor and the tracing that the cardiotocograph produces in response to the detected signals.

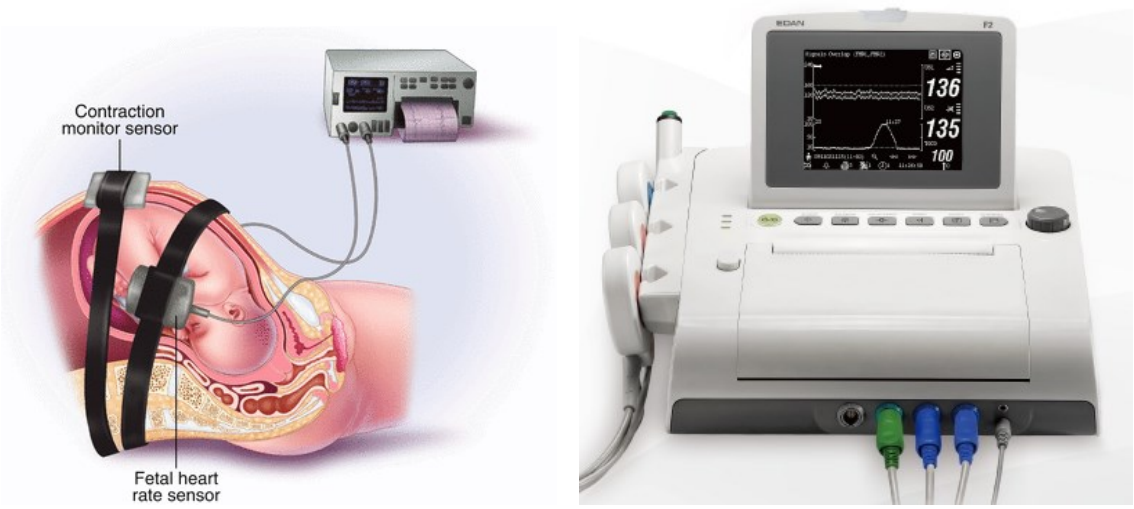

Fig. 5: cardiotocograph applied on the womb (left) and monitor (right).

The values of FHR and UC are shown, then, instant by instant on the cardiotocograph monitor recorded on graph paper similar to that used for the electrocardiogram (ECG). The resulting paper plot consists of two graphs, one below the other, representing on graph paper the two trends over time of FHR (top) and UC (bottom). Here is an example in Fig. 6.

The horizontal scale for recording CTG is called “paper speed” and can be 1, 2 or 3cm/min; in particular, in European countries, the use of a scale of 1cm/min is preferred. The vertical scale, on the other hand, can be 20 bpm/cm or 30 bpm/cm.

The FHR is recorded in bpm and can range from a minimum of 50 bpm to a maximum of 210-250 bpm; between the horizontal lines the distance is 5 bpm. The baseline is 130 bpm. The graph below, on the other hand, records UC in millimeters of mercury [mmHg]. They vary from a minimum of 0 mmHg to a maximum of 100 mmHg, and the lines are 10 mmHg apart. [16]

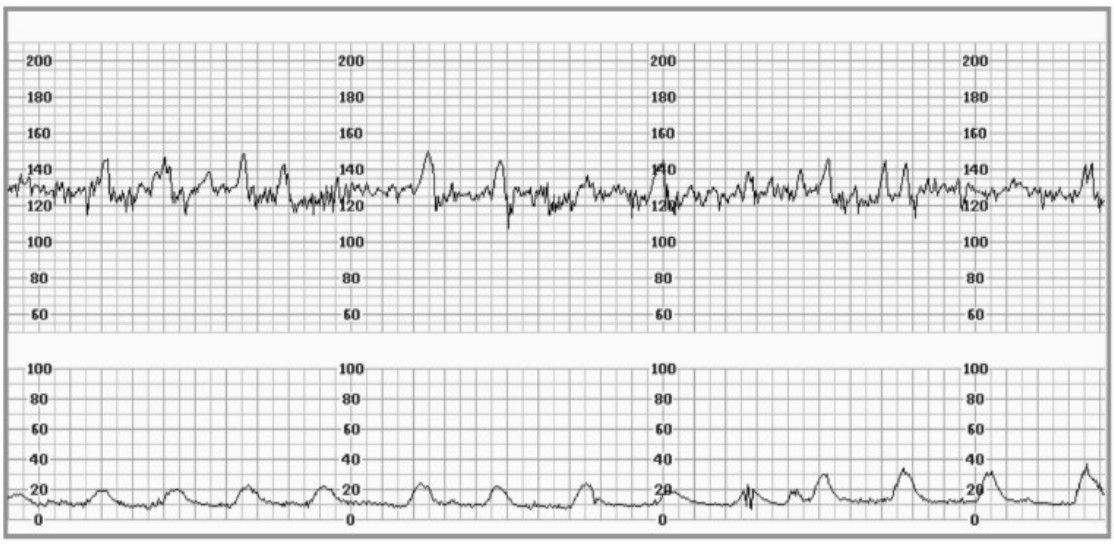

Fig. 6: FHR and UC monitoring, CTG trace.

2.2 Signals detected in CTG

As mentioned in the previous paragraph there are mainly two types of signals detected by CTG and they are as follows: the **FHR** and **Ucs**.

During external CTG, the patient wears a belt around the abdomen during monitoring so that movements are minimized. Another type of CTG is what is referred to as internal CTG. It is an alternative and very invasive as well as dangerous method that involves placing the electrode directly on the child's head. This form of continuous monitoring requires the rupture of the amniotic fluid and, again, a restriction of the mobility of the mother. Illustration of external and internal CTG in Fig. 7.

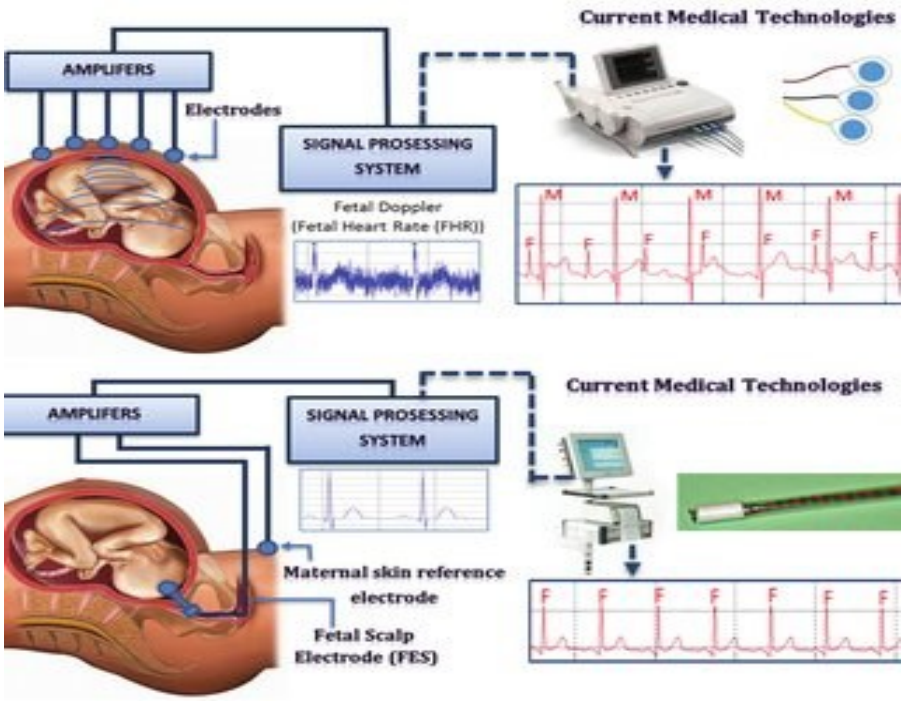

Fig. 7: external CTG (on the top) and internal CTG (on the bottom).

Both signals can be detected with either type of cardiotocographic examination.

- **Fetal Heart Rate**

- External monitoring: uses an ultrasound Doppler transducer to detect cardiac motion (both fetal heart rate and the speed and direction of blood flow in the fetal circulation). The resulting signal requires modulation and autocorrelation so that it has a quality suitable for storage. This process results in approximation of actual heart rate intervals but at the same time is considered accurate enough for subsequent analysis. This

method is more prone to signal loss, accidental recording of the mother's heartbeat, and signal artifacts. It also may not accurately record any fetal arrhythmias in some cases.

- Internal monitoring: instead uses a fetal electrode (usually placed on the fetal scalp), which assesses the time intervals between one heartbeat and the next by going on to identify certain characteristic features of the signal electrocardiographic ECG, which reflects the continuous change of potentials of action in correlation with the cardiac cycle. The external monitoring method thus allows the identification of the R peaks and the QRS complex, thus the entire ventricular depolarization cycle. This method allows a more accurate assessment of the intervals between cardiac cycles, but it is more expensive as it requires nonreusable electrodes as well as invasive. Internal monitoring causes perforated membranes and may induce contraindications mainly related to the risk of infection transmission [17].

- **Uterine Contractions**

- External monitoring: a tocodynamometer is used that assesses the increase in tension, measured through the abdominal walls. Improper positioning could reduce the tension applied to the supportive elastic, or abdominal adiposity could lead to incorrect registration of contractions. In addition, this technology provides accurate information only about the frequency of contractions and not their intensity or duration.

- Internal monitoring: instead, uterine contractions are measured using an intrauterine catheter that provides various information about contractions (frequency, duration, and intensity) and uterine tone. In contrast, this method is more expensive and may have contraindications, such as uterine bleeding from causes unknown, placenta previa, risks of fetal trauma, placental hemorrhage, perforation of the uterus, and infection. Repeated use of catheters for pressure intrauterine does not appear to be associated with improved outcomes in induced and enhanced labor, so it is not recommended for routine clinical use routine. [17]

Finally, maternal heart rate (MHR) can also be monitored simultaneously in particular cases when it becomes difficult to distinguish between fetal and maternal heartbeats.

2.3 Classification systems

There are different types of classification of cardiotocographic tracings that are enclosed in different guidelines. These guidelines are used to investigate and comprehend whether or not the fetus is in a potentially hazardous situation (shown by a suspicious tracing).

The parameters to be considered in the evaluation of CTG basically consist of:

- baseline
- variability
- accelerations
- decelerations
- contractile activity.

Baseline: represents the average rounded FHR in increments of 5 bpm during a 10-minute period, excluding marked variations in FHR corresponding to more than 25 bpm. It is represented graphically as a straight line that runs medially through the tracing, free of acceleration or deceleration. The FHR has a normal baseline if values range between 110 bpm and 150 bpm. In Fig. 8 it is shown an example of normal FHR between the correct range.

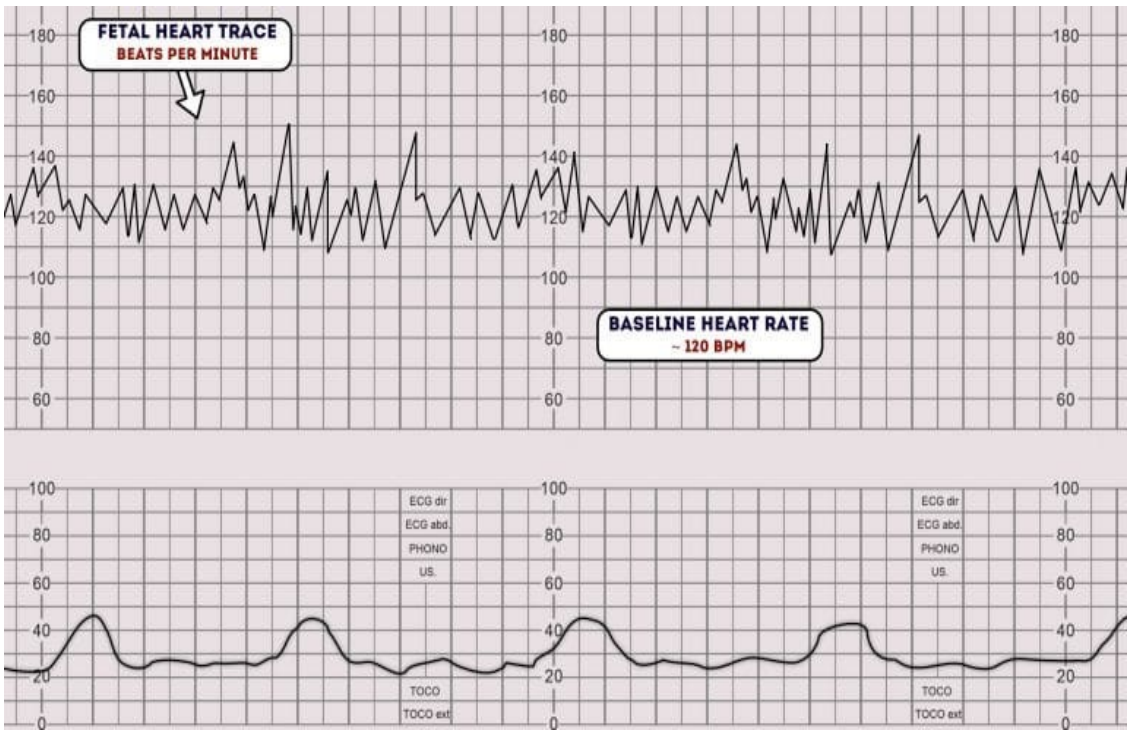

Fig. 8: representation of baseline in CTG trace.

Two abnormal conditions can be identified: tachycardia and bradycardia.

Tachycardia: the baseline has values around 160 bpm for more than 10 minutes. It can be caused mainly by fetal hypoxia, chorioamnionitis, hyperthyroidism, fetal or maternal anaemia, fetal tachyarrhythmia, hemorrhage and other causes. Fetal pathologies that can cause fetal tachycardia include: infections, prematurity of the central nervous system and anemia. It can also be asymptomatic. When the FHR is maintained consistently above 180 bpm constitutes a case of severe tachycardia. An example of severe tachycardia is shown in Fig. 9.

Bradycardia: in this case the baseline has values below 110 bpm for more than 10 minutes. Values between 100-110 bpm may be normal in the fetus, especially in post-dated pregnancies. It is considered severe bradycardia if the values are maintained below 100bpm. Maternal hypothermia, administration of certain agents and fetal arrhythmias such as atrioventricular blocks are possible causes. Other causes of severe and prolonged bradycardia can be prolonged cord compression, cord prolapse, epidural and spinal anaesthesia or maternal seizures.

An example of severe bradycardia is shown in Fig. 10.

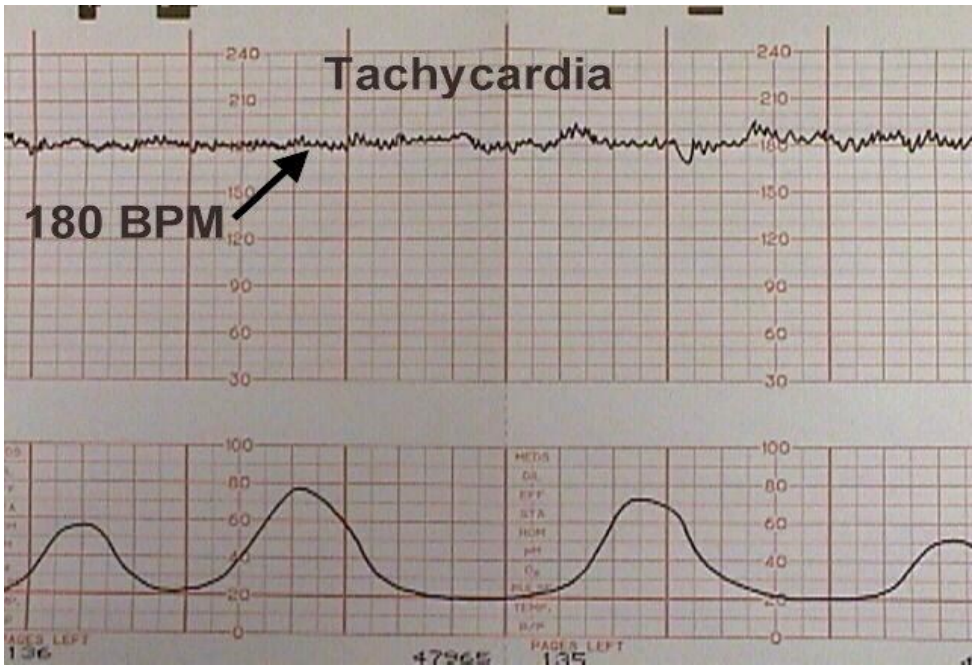

Fig. 9: example of severe tachycardia with 180 bpm.

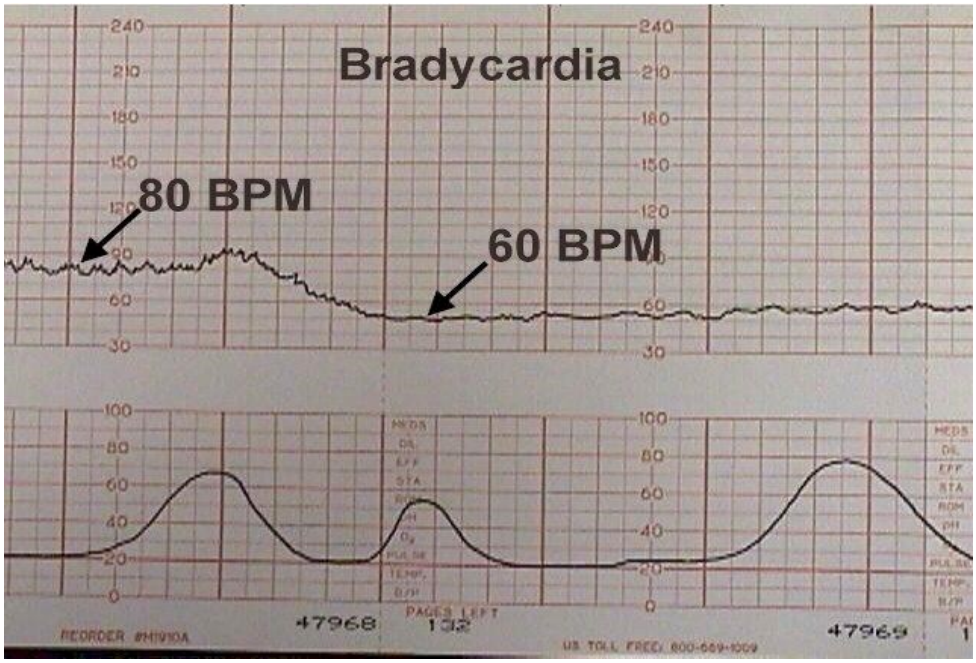

Fig. 10: example of severe bradycardia with less than 100 bpm.

Variability: refers to the oscillations of the FHR signal, measuring the difference between the maximum and minimum frequency reached in 1 minute of tracing. By norm, variability should be around 5 to 25 bpm. This phenomenon arises as a consequence of the interplay that takes place between the neurological system, chemoreceptors, baroreceptors, and the responsiveness of the heart. In point of fact, baroreceptors react to changes in arterial pressure, which triggers the activation of the sympathetic nervous system and causes vasoconstriction. On the other hand, chemoreceptors react to changes in the blood's partial oxygen pressure, which triggers the activation of the parasympathetic nervous system and causes bradycardia.

Because a healthy fetus will continually adjust its heart rate in reaction to changes in its surroundings, variability is a strong measure of how healthy a fetus is at that specific moment in time. For the purpose of calculating variability, it is necessary to determine the degree to which the highs and lows of the heart rate depart from the baseline rate (in bpm).

The CTG tracing in relation to variability may be classified as reassuring, non-reassuring, or abnormal, depending on the circumstances.

<u>Reassuring</u>: 5 – 25 bpm

<u>Non-reassuring</u>:

- less than 5 bpm for between 30-50 minutes

- more than 25 bpm for 15-25 minutes

<u>Abnormal</u>:

- less than 5 bpm for more than 50 minutes
- more than 25 bpm for more than 25 minutes
- sinusoidal behaviour.

Example of normal variability in Fig. 11.

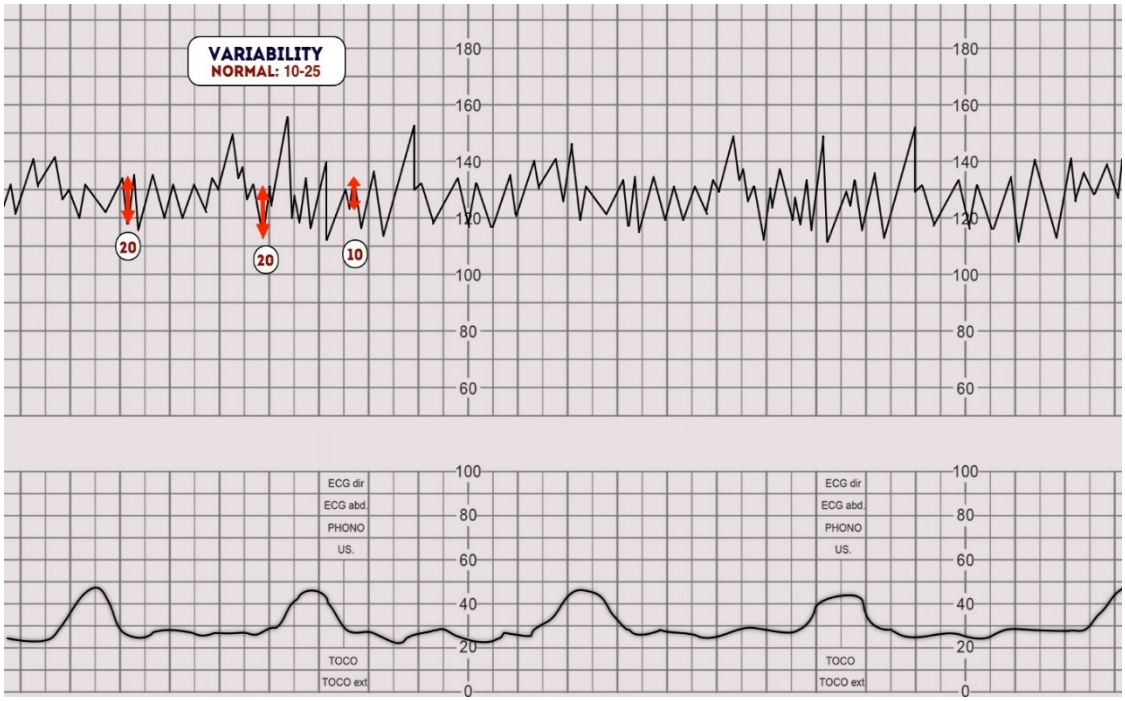

Fig. 11: variability normal between 10 and 25 bpm.

Also, can be identified two pathological conditions: reduced and increasing variability.

Reduced variability: it is a band length below 5bpm for more than 50 minutes, or for more than 3 minutes during decelerations. Among the main causes we find central nervous system hypoxia, acidosis, prematurity, fetal sleeping, tachycardia or neurological abnormalities (Fig. 12).

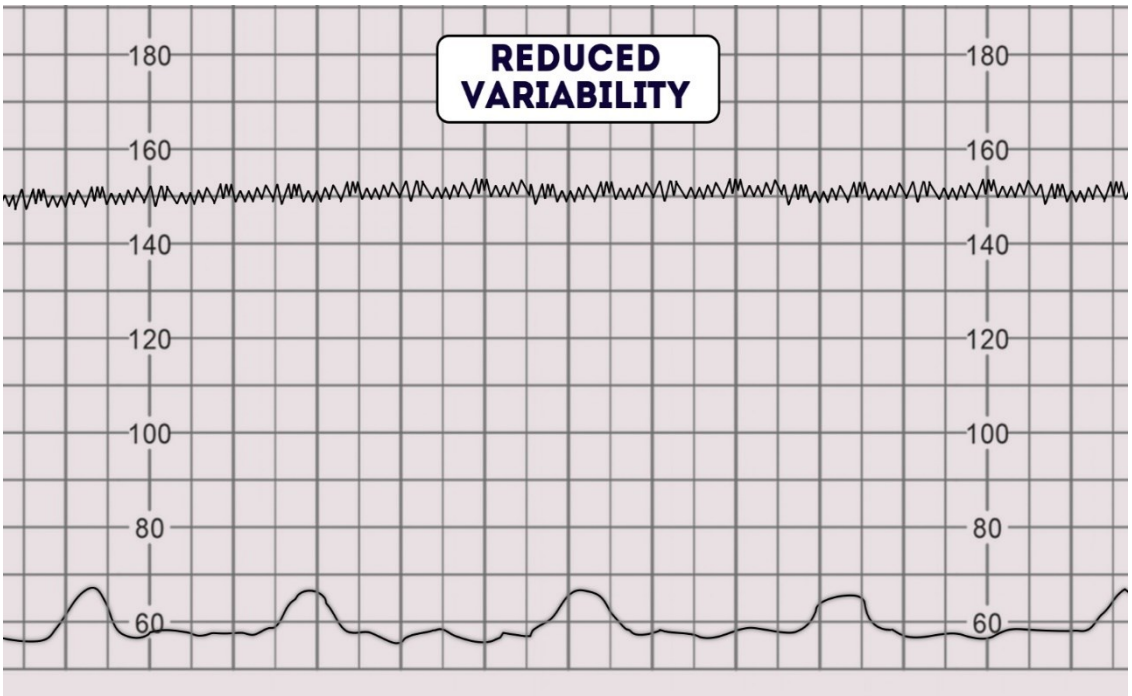

Fig. 12: Example of reduced variability of bpm.

Increasing variability: when the bandwidth is above 25bpm for more than 30 minutes and the tracing has a jumping appearance. The pathophysiology of this field is not yet fully understood, but it can be recognized in recurrent decelerations when hypoxia or acidosis evolve very rapidly. It is hypothesized to be the cause of fetal autonomic instability or hyperactivity by excessive activity of the parasympathetic system: stimulation of the vagal system by head compression.

All the changes in the rate of the FHR in response to stimulations. Accelerations or decelerations of the heartbeat can be seen.

Accelerations: are sudden increases in FHR relative to baseline of more than 15 bpm in height and from the duration of more than 15 seconds but for less than 10 minutes. Below 32 weeks of gestation the increase of 10bpm for 10 seconds is considered normal. Most of the accelerations coincide with fetal movements (especially if they are sporadic) and are a sign of a fetal neurological response that does not involve hypoxia or acidosis. If they are periodic, they could be related both to movements fetal but also to uterine activity. Before 32 weeks of gestation, their amplitude and frequency should be lower. After 32 to 34 weeks, accelerations are rare in periods of deep sleep, in which they can, however, last up to 50 minutes. If such accelerations coincide with the Ucs, is frequent to consider a misrecording of the FHR, since it is more common for the FHR to decelerate

with a contraction while the maternal heart rate is increasing. In principle, however, it can also be considered a sign of fetal health (Fig. 13).

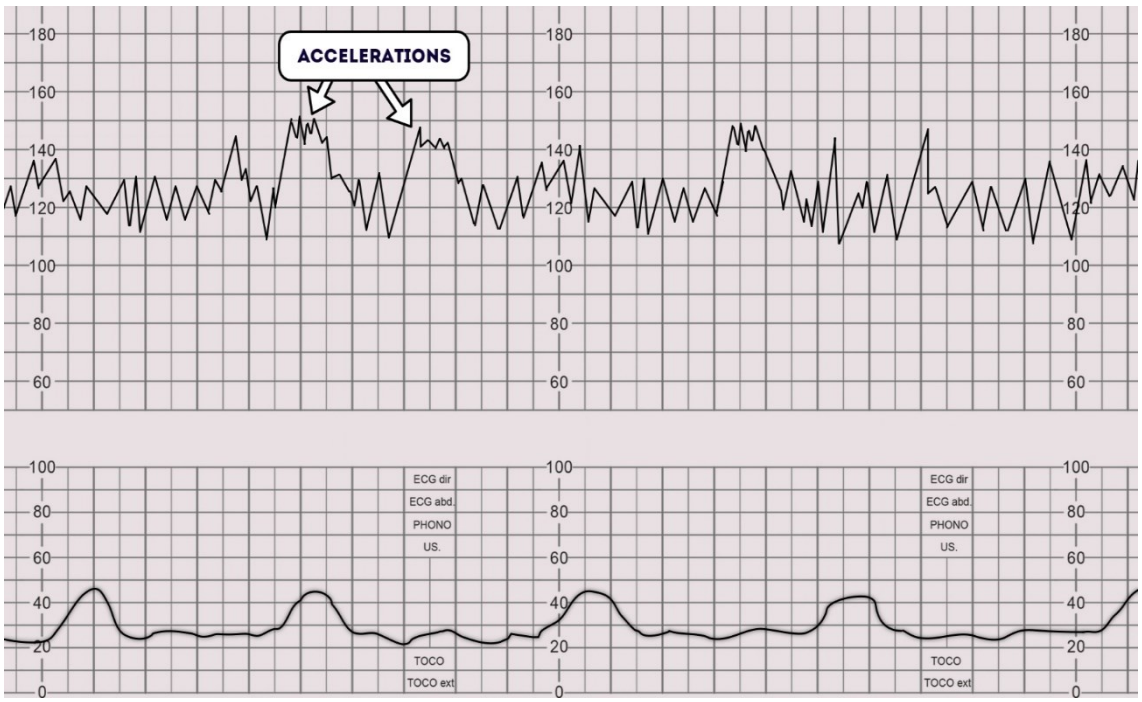

Fig. 13: accelerations.

Decelerations: are moments in the CTG tracing at which there is a decrease in FHR with subsequent return to baseline, for a duration of more than 15 seconds. The autonomic and somatic nervous systems work together to regulate the rate at which the fetal heart beats. As a defence mechanism against hypoxic stress, the developing fetus slows its heart rate in order to maintain adequate oxygenation and perfusion of the myocardium. In contrast to an adult, a fetus is unable to increase the depth or pace of its respiration. A deceleration refers to the slowing down of one's heart rate for the purpose of lowering the demand placed on the myocardium.

There are a variety of distinct sorts of decelerations, each of which carries with it a unique set of implications. In relation to the morphology, decelerations are classified, by relationship to contraction and by morphology.

- Early decelerations: these are weak, of short duration and with a variability normal. Start when the Ucs begins and recover when these latter stops**.** They are characterized by reaching nadir (lowest point reached by the deceleration) in a time less than or equal to 30 seconds. They are thought to be caused by compression of the fetal head and do not indicate hypoxia or fetal acidosis (Fig. 14).

They are considered to be physiological and not pathological.

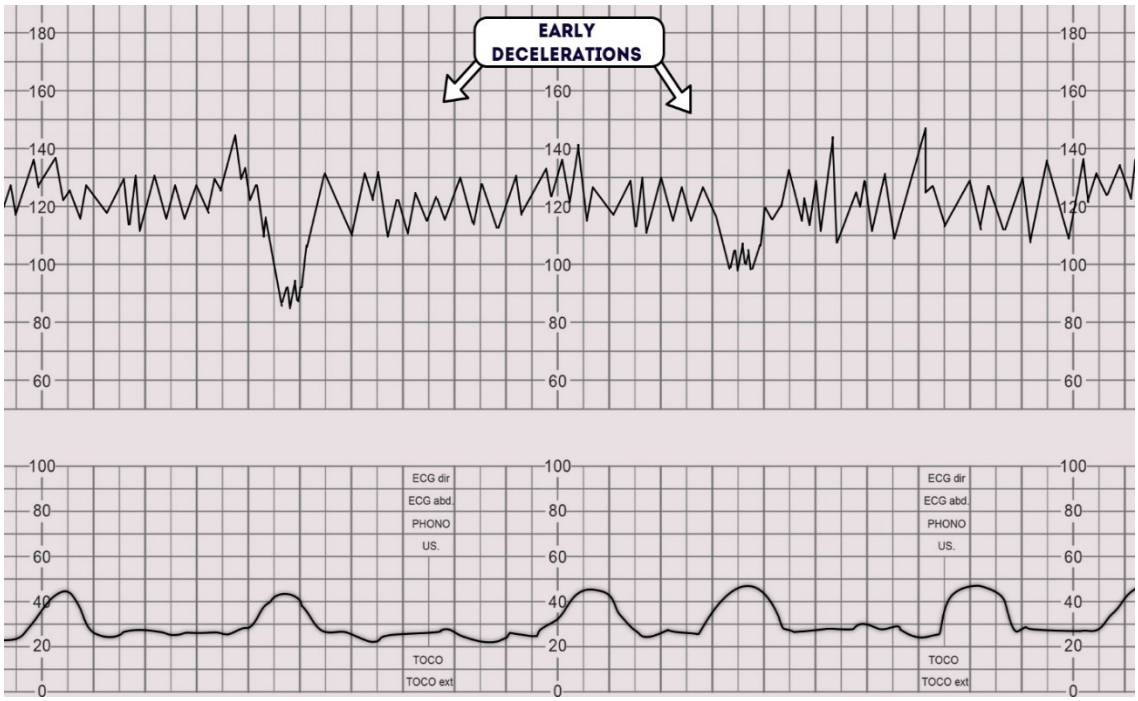

Fig. 14: early decelerations.

- Variable decelerations: are defined as such because they can occur in random relationship with contraction but variable in its repetitiveness and with variable morphology. They show rapid lowering and then decreasing, good variability without the decelerations, a rapid return to baseline, variations in size and shape and relationships with uterine contractions. They are abrupt decreases in fetal rate with variable morphology, characterized by at least 15 beats below baseline for longer than 15 seconds and less than 2 minutes. The presence of intermittent variable decelerations (<50% of Ucs) often requires no treatments. If they are recurrent (>50% of Ucs), they require intrauterine resuscitation techniques.

Variable decelerations are usually caused by umbilical cord compression: the umbilical vein is occluded causing an acceleration of FHR in reaction. Then the umbilical artery is occluded provoking a rapid deceleration. Then the pressure on the cord is reduced other acceleration happens and the baseline rate returns (Fig. 15), [18].

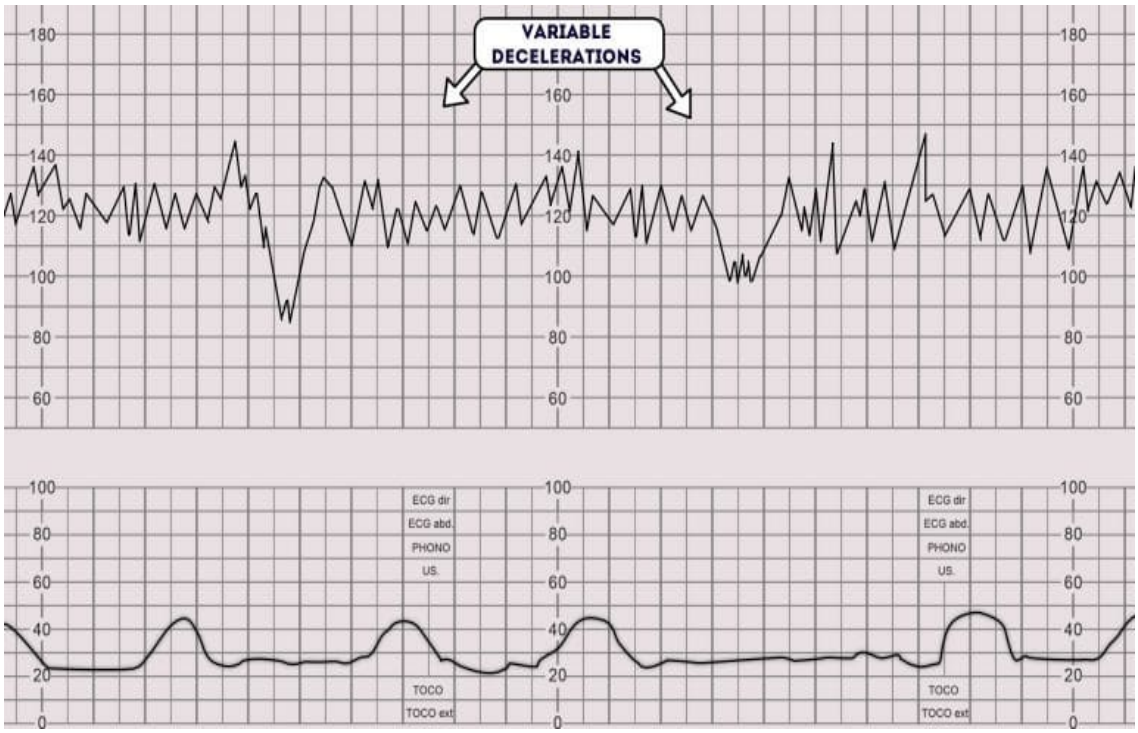

Fig 15: variable decelerations.

- Late decelerations: these are decelerations with a gradual onset and/or a return gradual to baseline and/or reduced variability during deceleration. A gradual onset and end occur when more than 30 seconds elapse between the beginning and end of the deceleration. When contractions are monitored adequately, we notice that such decelerations begin after more than 30 seconds from the start of the contraction and there is a return to baseline after the end of the contraction itself. The presence of recurrent late decelerations is thought to may reflect transient or chronic utero-placental insufficiency. The most common causes include maternal hypotension, pre-eclampsia and fetal hypoxia (Fig. 16).

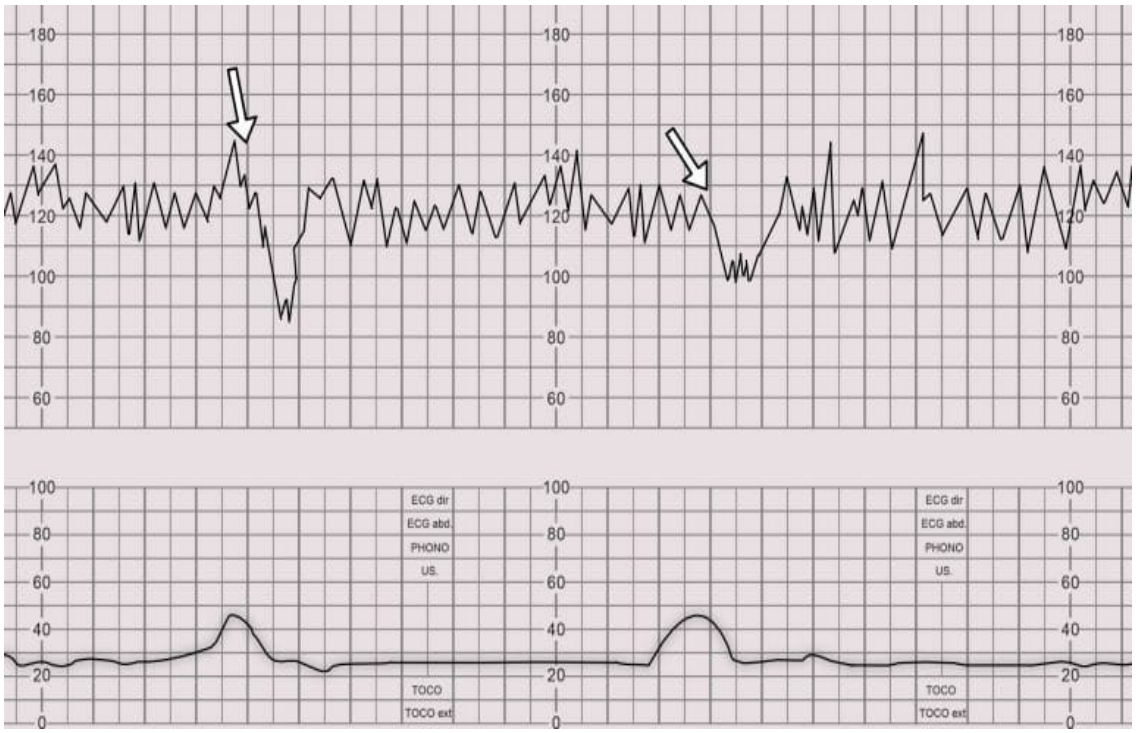

Fig. 16: late decelerations.

Decelerations are uniform when they are characterized by constant parameters. Their duration ranging from 20 to 90 seconds, with an amplitude of about 50 bpm. Their decrease is approximately 20 to 30 bpm with a frequency greater than 100 bpm.

- Prolonged decelerations: this is a decrease in FHR for more than minutes. A longer duration with a signal that remains below 80bpm and reduced variability without decelerations may be a symptom of acute hypoxia or fetal acidosis and require medical intervention (Fig. 17).

If it lasts between 2-3 minutes it is classified as non-reassuring.

If it lasts longer than 3 minutes it is immediately classified as abnormal.

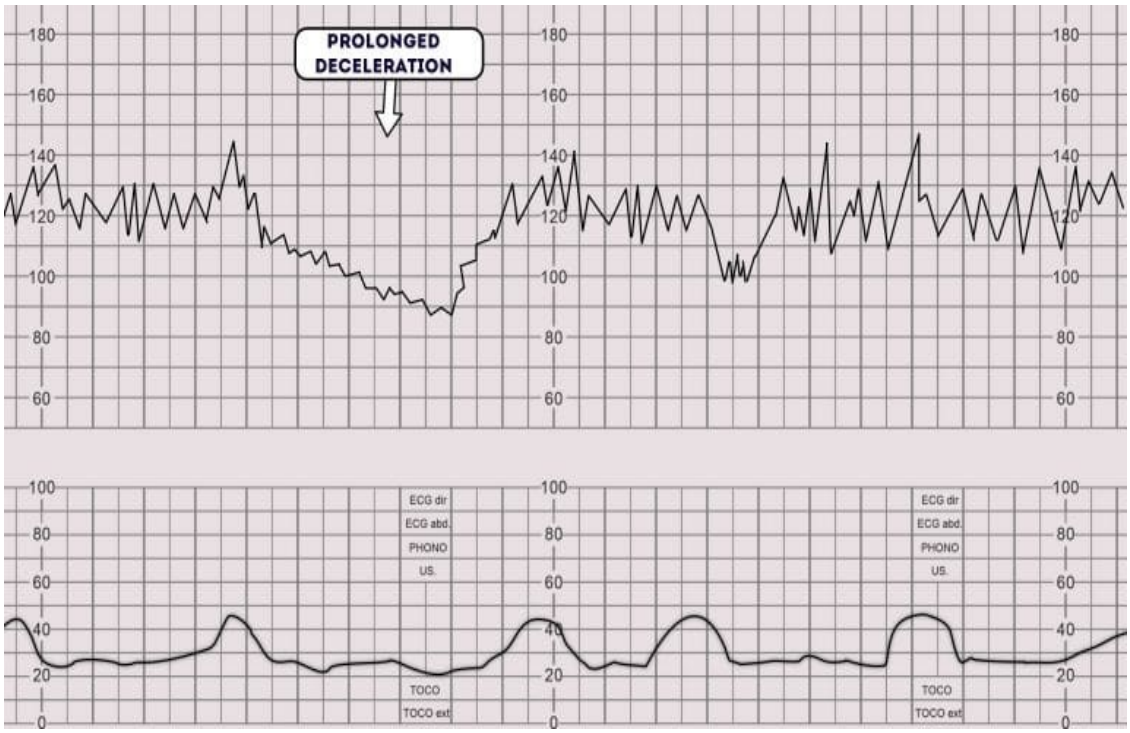

Fig. 17: prolonged deceleration.

- Sinusoidal pattern: CTGs with a sinusoidal pattern are rare but of great concern as they have very high fetal mortality rates [18]. A sinusoidal pattern usually indicates severe fetal hypoxia, severe fetal anemia, and fetal or maternal hemorrhage. Is characterized by a regular and uniform wave-like pattern with a frequency of about 2-5 cycles per minute, base frequency around 120-160 bpm and no beat-to-beat variability (Fig.18).

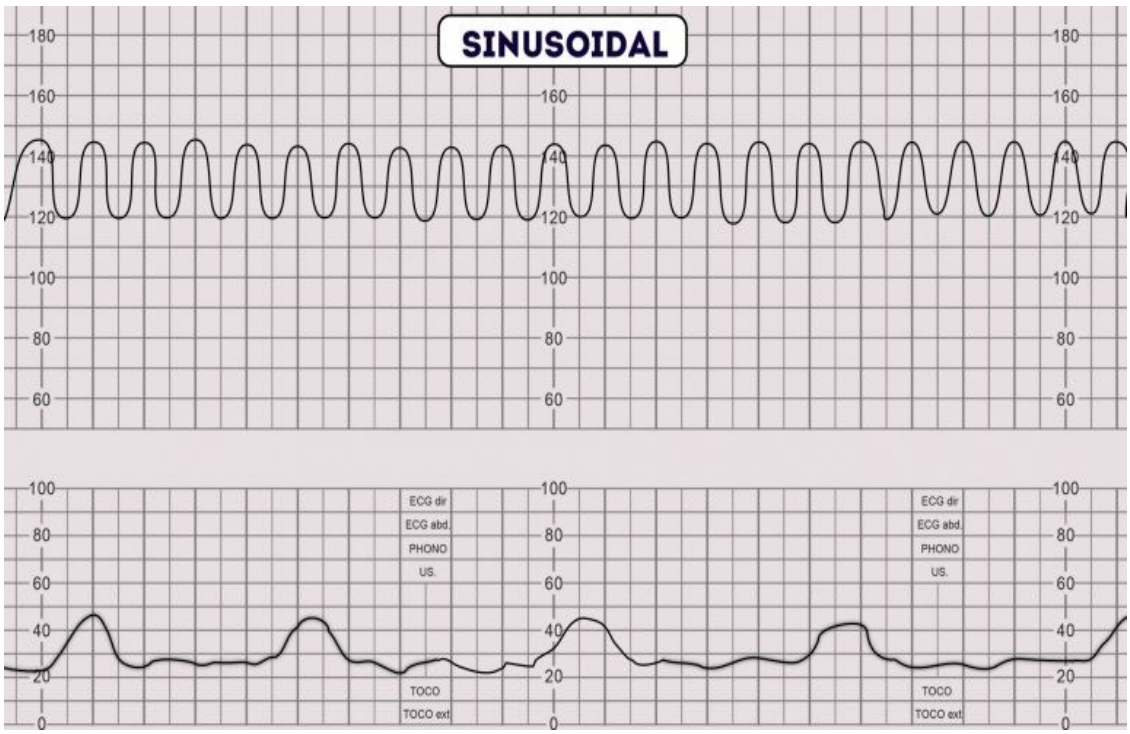

Fig. 18: Sinusoidal CTG.

Contractile activity: leads to a gradual “bell-like” increase in the uterine activity signal followed by an almost symmetrical decrease, with a duration of 45-120 seconds. With the tocodynamometer, it is possible to detect in a reliably only the frequency of contractions, while intensity and duration can be assessed from the FHR graph. A phenomenon called tachysystole can be encountered: it occurs when there is a high frequency of contractions or when there are more than 5 contractions in 10 minutes or on average in each 30-minute period.

2.4 The main guidelines for CTG tracings

Since the introduction of CTG, many different definitions and guidelines have been created for the purpose of creating a classification of the five parameters analyzed in CTG tracings, which are baseline, variability, accelerations, decelerations, and contractile activity. After nearly 30 years, the 1987 International Federation of Gynecology and Obstetrics (FIGO) guidelines still remain the only international consensus document that many national guidelines refer to. Many countries have adopted some sort of classification system for CTG patterns. Most guidelines provide a three-level classification, but with some differences in divisions and nomenclatures, as shown in Table 2. All classifications are non-evidence-based documents.

The main classifications for CTG patterns include the following:

- **FIGO**: International Federation of Gynecology and Obstetrics
- **SFOG**: Swedish Society for Obstetricians and Gynecologists
- **NICE**: National Institute of Health and Care Excellence
- **ACOG**: American Congress of Obstetricians and Gynecologists
- **RANZCOG**: The Royal Australian and New Zealand College of Obstetricians and Gynaecologists
- **SOGC**: Society of Obstetricians and Gynaecologists of Canada.

FIGO	SFOG	RCOG/NICE	ACOG	RANZCOG	SOGC
Normal	**Normal**	**Reassuring**	**Category I**	**Normal**	**Normal**
FHR 110-150 bpm Variability 5-25 bpm	FHR 110-150 bpm Variability 5-25 bpm ≥ 2 accelerations / 60 min Early decelerations Variable uncomplicated decelerations with duration < 30 sec and amplitude < 60 bpm	FHR 110-160 bpm, Variability ≥ 5 bpm, no decelerations, present accelerations	FHR 110-160 bpm, Moderate variability, Absent late or variable decelerations, Early decelerations and accelerations can be present or absent	FHR 110-160 bpm, Variability 5-25 bpm Accelerations present No decelerations	FHR 110-160 bpm, Variability 6-25 bpm, ≤ 5 bpm for < 40 min, None or occasional uncomplicated variable or early decelerations, Accelerations spontaneously or present with fetal scalp stimulation
Suspicious	**Aberrant**	**Non-reassuring**	**Category II**	**Abnormal**	**Atypical**
FHR 150-170 bpm or 100-110 bpm Variability 5-10 bpm >40 min Variability > 25 bpm Variable decelerations	FHR 100-110 bpm FHR 150-170 bpm FHR < 100 bpm ≥ 3min Variability < 5 bpm > 40 min without accelerations Increased variability < 2 accelerations/60 min Variable uncomplicated decelerations lasting 30-60 sec and / or amplitude > 60 sec (A combination of ≥ 2 aberrant parameters makes the tracing suspiciously pathological)	FHR 100-109 bpm FHR 161-180 bpm Variability < 5 bpm 40-90 minutes, Typical variable decelerations more than 50% of contractions > 90 min. Single prolonged deceleration up to 3 min (Absence of accelerations with otherwise normal trace is of uncertain significance)	All FHR tracings not categorized as I or III.	*but unlikely associated with significant fetal compromise* FHR 100-109 bpm or Absence of accelerations Early decelerations Uncomplicated variable decelerations	FHR 100-110 bpm FHR >160 bpm for 30-80 min Variability ≤ 5 bpm for 40-80 min Repetitive (≥3) uncomplicated variable decelerations Occasional late decelerations Single prolonged deceleration 2-3 min, Absence of acceleration with fetal scalp stimulation
Pathological	**Pathological**	**Abnormal**		**Abnormal**	**Abnormal**
FHR < 100bpm or > 170 bpm Variability < 5 bpm > 40 min Severe variable decelerations or severe repetitive early decelerations Prolonged decelerations Late decelerations Sinusoidal pattern	FHR > 170 bpm FHR < 100 bpm ≥ 3 min Variability < 5 bpm > 60 min without accelerations Sinusoidal pattern Variable complicated decelerations with duration > 60 sec Uniform late decelerations Combined decelerations	FHR < 100 bpm FHR >180 bpm Sinusoidal pattern ≥ 10 min, variability < 5 for ≥ 90 min. Either atypical variable decelerations over 50% of contractions or late decelerations, both for > 30 min. Single prolonged deceleration > 3 min.		*and may be associated with significant fetal compromise* Tachycardia Reduced variability Complicated variable decelerations Late decelerations Prolonged decelerations	FHR <100 bpm FHR>160 for > 80 min Erratic baseline Variability ≤ 5 bpm > 80 min or ≥ 25 > 10 min. Sinusoidal pattern ≥3 complicated variable decelerations Late decelerations > 50% of contractions Single prolonged deceleration 3-10 minutes Absent or present accelerations
	Preterminal		**Category III**	**Abnormal**	
	Absent variability without accelerations, regardless decelerations and baseline FHR		Absent variability and either: Recurrent late decelerations, recurrent variable decelerations, or bradycardia Sinusoidal pattern	*and very likely to be associated with sign fetal compromise* FHR < 100 bpm > 5min Absent variability Sinusoidal pattern Reduced variability with complicated variable or late decelerations	

Table 2: main guidelines for CTG tracings.

As can be seen from Table 2, the guidelines have differences in classification regarding parameters and in some cases even terminology. Indeed, it can be seen that for NICE, the categories are divided into “reassuring,” “not reassuring,” and “abnormal.” AGOC, on the other hand, is the only guideline that divides the level of fetal health hazard into “categories” (from one to three expressed in Roman numerals). For FIGO and SFOG, the third level is named “pathological” while for the rest except ACOG, it is named “abnormal.” The second level is different among all guidelines. The fourth level exists only for SFOG, ACOG and RANZCOG. The differences in terminology at any rate do not create serious problems in interpreting the tracks, usually one chooses to follow one of them and acts accordingly. On the other hand, as far as classification parameters are concerned, in some cases some turn out to be more accurate than others.

In any case, a key role is handled by the physicians involved in the interpretation of the tracings. The risk of misdiagnosis can lead to serious consequences. Hence, the guidelines are the starting point for interpretation, the physician’s preparation and experience also play an important role, but they do not ensure the correct diagnosis.

CHAPTER THREE: UNSUPERVISED DEEP-LEARNING METHODS

3.1 Introduction

Deep learning is a subcategory of machine learning and denotes that branch of artificial intelligence (AI) that refers to algorithms inspired by the structure and function of the brain called deep neural networks (DNN), (Fig. 19).

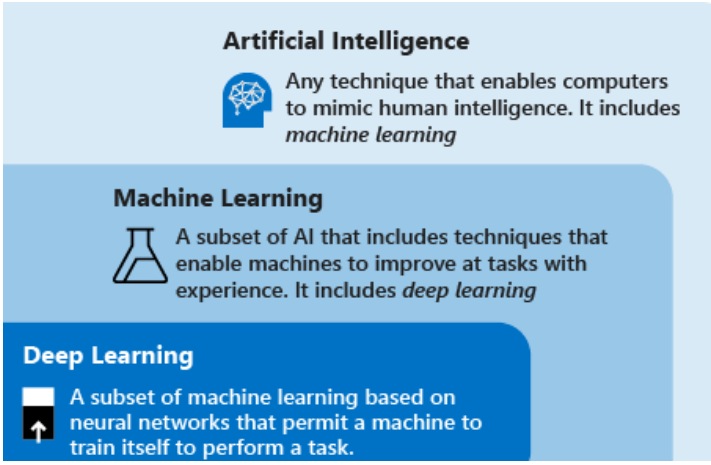

Fig. 19: illustration of artificial intelligence, machine learning and deep learning.

Deep learning is founded on the concept of artificial neural networks (ANN). Because the structure of ANN is composed of numerous layers of input, output, and hidden connections, the learning process is a complex one. Each layer of the model comprises units that are responsible for transforming the incoming data into information that the following layer may utilize for a certain type of prediction activity.

Machine learning is a subfield of AI that makes use of methods (such as deep learning) that enable computers to learn from their mistakes and become more efficient as they complete jobs. The following stages constitute the various components of the learning process:

- Input data into an algorithm: during this stage of development, more data may be input into the model in a variety of ways, including through the use of feature extraction.
- Make use of these data to train a model using the appropriate method.
- Perform validation and deployment of the model.

- Utilize the decentralized approach in order to carry out a process that is automatically predictive.

In general AI studies the ways in which machines can mimic human intellect. Machine learning is a component of it. [19]

In fact, machine learning and deep learning both describe methods for teaching computers to learn and make decisions.

Some important divergences between machine learning and deep learning make them suitable for different applications. Classical machine learning often includes feature engineering by programmers that helps the algorithm make accurate predictions on a small data set. Deep learning algorithms are generally designed with more layers of decision making to require less specific feature engineering. Also, deep learning is traditionally used for larger data sets so that networks or algorithms can be trained to make many layered decisions.

This chapter will discuss the different types of learning with special attention devoted to deep learning and its network topologies.

3.2 Deep Learning

Firstly, a brief description of the origins and functioning of deep learning in order to better comprehend its nature and its architecture.

In 1986, Rina Dechter coined the phrase “deep learning”. Making an intelligent computer that emulated the human brain was the primary driving force behind the development of the discipline of deep learning [20].

The brain is the most significant and decision-making organ in humans because it makes choices depending on what it sees, smells, feels, and hears. The brain can also retain memories and use its own experiences to tackle complicated issues.

Researchers have hoped to create a computer that is as clever as our brain in recent years by researching the biological makeup and functioning of the brain.

The word “deep” in deep learning refers to the number of layers through which data is transformed from input to the desired output, as was indicated in the previous section.

Any form of computer software that can learn on its own without being specifically written by the programmer is known as a machine learning program. [20]

The first neural network-based computer model was developed in 1943 by Warren McCulloch and Walter Pitts, which marked the beginning of deep learning. The main objective was to simulate how the human brain thinks.

Alan Turing is considered the father of AI.

The first single- and multilayer ANNs were introduced by Rosenblatt in 1957.

Even though the deep learning group has made significant contributions, just a handful of its major proponents (Yann LeCun, Geoffrey Hinton, and Yoshua Bengio) were awarded the Turing Prize in 2018.

An eight-layer deep network developed using the group approach of data processing algorithm was first reported in 1971.

Between 1970 and 1986, the concepts of backpropagation, recurrent neural networks (RNNs), and limited Boltzmann machines (BM) were introduced.

Convolutional neural networks (CNN), bidirectional RNNs, and long short-term memories (LSTM) were the state of the art between 1979 and 1998.

Geoff Hinton first presented the deep belief network (DBN) in 2006 [20].

The first distinction that can be made about deep learning is certainly in the type, in fact unsupervised learning and supervised learning are the two main categories of learning and will be further discussed in the next paragraph.

Supervised learning, use completely labelled data to instruct or train the computer. The computer then draws knowledge from the labelled data and prepares itself for unexpected input. In supervised learning, the machine can only provide accurate results when the input has previously been used during the training phase. This means that the more experience the machine has, or the larger its training dataset, the more likely it is to produce accurate results.

As opposed to supervised learning, *unsupervised learning* allows a model to function independently by collecting more data and finding information within the data. Unsupervised learning is more difficult since it frequently works with unlabelled data [20].

3.3 Learning types

As was discussed in the preceding paragraph, learning algorithms may be segmented into a variety of categories, each of which is determined by the end result that the algorithm is intended to produce. Supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning are the four most prevalent kinds of algorithms [21]. Unsupervised learning with its application is the subject of investigation in this thesis.

3.3.1 Supervised learning

It is regarded to be the most essential approach of machine learning, which is the capacity of computers to deduce a function of various tasks based on supervised training data. This ability is what gives machine learning its name. In most cases, training data is represented by a collection of training samples [22]. Every model comprises a pair that consists of input data, which is often a vector, and a desired outcome as an output value, which is also referred to as a supervised signal. In supervised learning, one learns a mapping between a collection of input variables and an output variable, and then uses this mapping to make predictions about data outputs that have not yet been seen. An ideal consequence is one that enables the algorithm to accurately establish class labels for examples that have not yet been encountered. The capacity to create adequate outputs for inputs that were not encountered during training is the primary characteristic of supervised classifiers [21]. This ability is often referred to as generalization ability.

Classification is the standard formulation of the supervised learning task, in which the learner is tasked to learn the behavior of a function that maps a vector to one of several classes by observing several input-output instances of the function [21]. The learner must learn this function by observing several input-output instances of the function.
During the building phase, the training operation will continue as long as necessary in order for the algorithm to become capable of achieving the highest possible accuracy on the data that is provided.

3.3.2 Unsupervised learning

Unsupervised learning is another form of machine learning model that may be used to discover implication from training datasets that include input data but do not involve output data. This type of learning does not need human supervision (unlabelled responses). In unsupervised learning, the system learns from a collection of distinct input patterns that represent statistical structure without a defined output aim or environmental judgment for each input [23]. This kind of learning is referred to as "learning by observation." Learning without supervision is more analogous to how the human brain works than the alternative, guided learning. For instance, the human eye contains approximately 10 million photoreceptors that are constantly changing their activity in response to the visualization of various objects. These photoreceptors provide all of the information that indicates what the objects are and, as a result, enable recognition of their shape and distance from the observer. Reflexes also enable a prompt response of the human body (as well as an adjustment to the body's postural alignment). It is well knowledge that the activity pattern of sensory neurons has an effect on the anatomical and physiological features of synapses found in the neocortex of the brain. Because of all this information, unsupervised learning is an absolutely necessary component of computational models for synaptic variation. In addition, the relevance of this kind of education cannot be overstated when it comes to classifying variables into previously undiscovered subgroups.

3.3.3 Semi-Supervised learning

The study of how computers and natural systems, such as people, learn in the presence of both labelled and unlabelled input is the focus of the learning paradigm known as semi-supervised learning, which examines this phenomenon. In comparison to the availability of large amounts of labelled data, the availability of large amounts of unlabelled data has motivated machine learners to investigate new methods that can use information about the distribution of inputs. Additionally, the labelling or annotation of data can be expensive [8].

Learning that is semi-supervised, also known as SSL, may be thought of as a compromise between learning that is supervised and learning that is unsupervised. The most common kind of semi-supervised learning divides the data set into two parts, with one

portion including labels and the other part containing points for which the labels are unknown [23].

In practice, semi-supervised learning is beneficial when there is a significant amount of unlabelled data relative to the amount of labelled data. Because it can use unlabelled data to improve supervised learning tasks, this type of algorithm is an important tool in machine learning. It can produce a significant improvement in accuracy, especially when data are expensive and sparse [24]. In addition, it can use unlabelled data to improve unsupervised learning tasks. This is the case, for instance, in a great number of different application areas of machine learning: implemented into speech recognition, capturing massive volumes of speech at nearly no cost is possible, but labelling it needs a person to listen to it and write down the transcript. There are billions of web pages that are immediately accessible for automated processing, but in order to accurately categorize them, it is necessary for people to read them.

3.3.4 Reinforcement learning

In the field of machine learning, the term “reinforcement learning,” abbreviated “RL,” refers to a group of approaches that enable an algorithm to learn from the regular effects of its actions rather than being explicitly taught based on the algorithm’s prior experiences. The machine is put into a setting where it may teach itself in an ongoing manner via the use of trial and error [25]. This computer learns from previous experiences and strives to collect the finest available information in order to make appropriate judgments about commercial matters.

The algorithm receives instructions from the reinforcement machine learning process on the policy of how to behave in response to an observation of the world. Every action has some kind of influence on the environment, and the input that the algorithm for learning receives from the environment helps to direct its development [26]. The focus of this kind of machine learning would not be on determining which action should be taken into consideration; rather, it would be on determining which action created such an exceptional reward. RL is often used in a variety of applications with the purpose of resolving a number of difficult issues. For instance, RL has proven successful in the fields of medical diagnosis, voice recognition, bioinformatics, computational vision, spell recognition, and robot locomotion [21].

3.4 Deep Network algorithms

Deep learning, much like machine learning, may be carried out using a variety of topologies and algorithms. However, deep learning can be carried out on large datasets considering the nature of learning that enables us to examine the additional division (supervised and unsupervised).

Listed below are the primary topologies for deep learning [20]. Each of these methods can be used in a supervised and unsupervised manner depending on whether they include pattern analysis (unsupervised) or classification analysis (supervised) [27].

3.4.1 Deep neural network (DNN)

Between the input and the output of a DNN system lies a multilayer perceptron, also known as a hidden layer. Because each layer is linked to the one that came before it, the network is able to make an accurate estimate of the output by traversing each layer in turn. This estimation is based on the weights and the activation function. We are able to model any kind of complicated non-linear connection using DNN. The ability of the DNN to acquire knowledge about the feature that is most important to the targets [28] serves as the network's primary support structure. Research holes may be filled in the DNN's model selection and training dynamics via the use of graph CNN combination optimization and Bayesian neural networks for uncertainty estimation. DNN is useful for a wide variety of purposes, including medical diagnosis, computer vision, machine translation, filtering for social networks, playing boards, and video games (Fig. 20), [20].

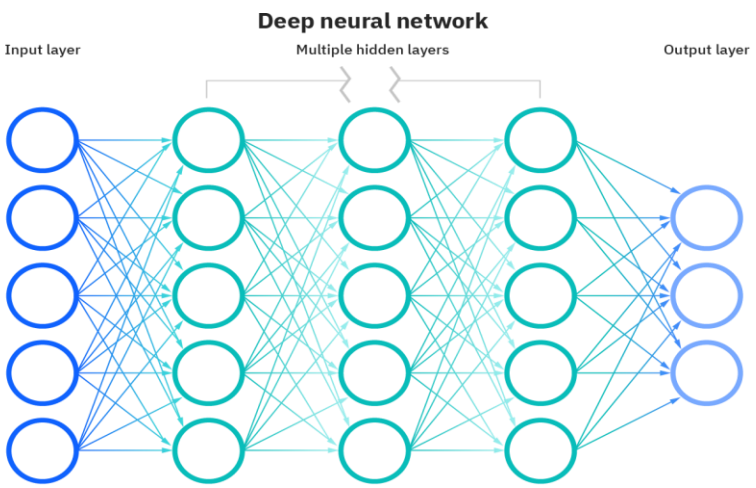

Fig. 20: Deep Neural Network.

3.4.2 Recurrent Neural network (RNN)

RNNs are a kind of deep learning network that are used particularly in situations in which there is sequential data or a time-series, such as in video, audio, and other similar applications. In most cases, the RNN preserved the data from the previous state all the way through to the next state. It is referred to as recurrent because it executes the same function for each input, but the result is always unique due to the fact that it also takes into account the results of earlier computations. The Long-Short-Term Memory Network (LSTM) is the cutting-edge research subject in the field of deep learning using RNN. RNN offers a solution to a wide variety of issues, including tackling time-varying matrix inversion [29], intelligent transportation system [30], and a great deal of additional issues. Evaluation of sentences and processing of linguistic data are two areas in which the RNN excels (Fig. 21), [20].

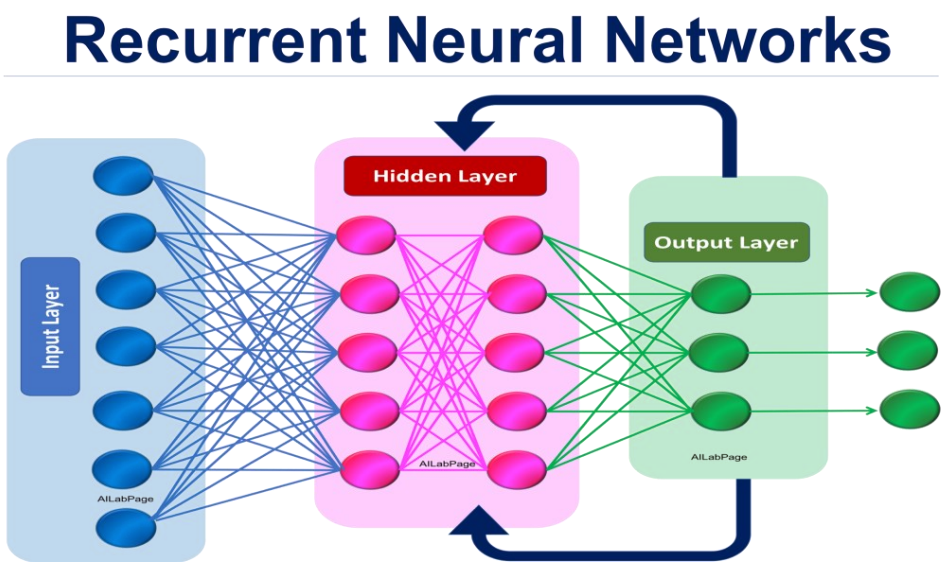

Fig. 21: Recurrent Neural Network.

3.4.3 Deep belief network (DBN)

The Deep Belief Network (DBN) is a technique for probabilistic unsupervised deep learning. It conceals many different elements behind its various layers. It requires more hidden layers in order to tackle the more difficult issues; each layer has a unique statistical link with the other levels. DBN is able to learn probabilistically, but BDN requires training under supervision after it has learned in order to accomplish classification. The DBN can identify clusters and produce photos, video sequences, and

motion-capture data. Additionally, the DBN can create images. In Fig. 22 an example of how DBN works in models for stress classification [20].

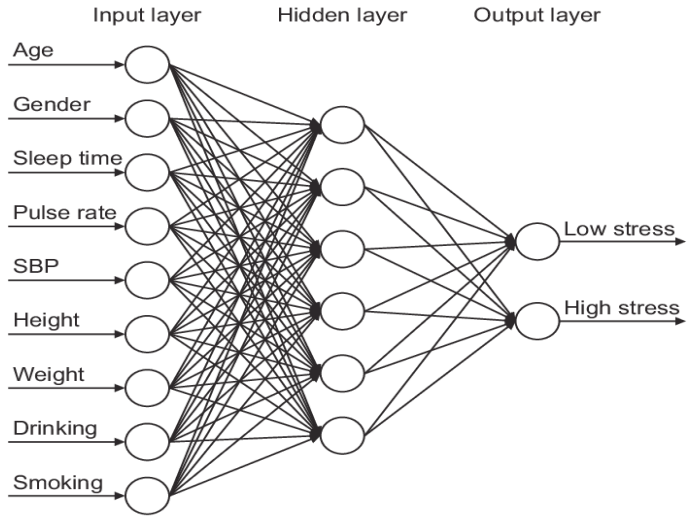

Fig. 22: Deep Belief Network.

3.4.4 Boltzmann machine (BM)

The BM is a network that is a neuron-like unit that is evenly connected. This neuron-like unit is responsible for making choices about whether or not to be off or on in a stochastic manner [20].

BM is used to find solutions to computational issues such as learning challenges, search problems, and optimization problems. The learning process reveals a great many characteristics, each of which exhibits quite complicated behaviour in the training dataset. The Boltzmann machine is used for the purposes of classification and the reduction of dimensionality (Fig. 23).

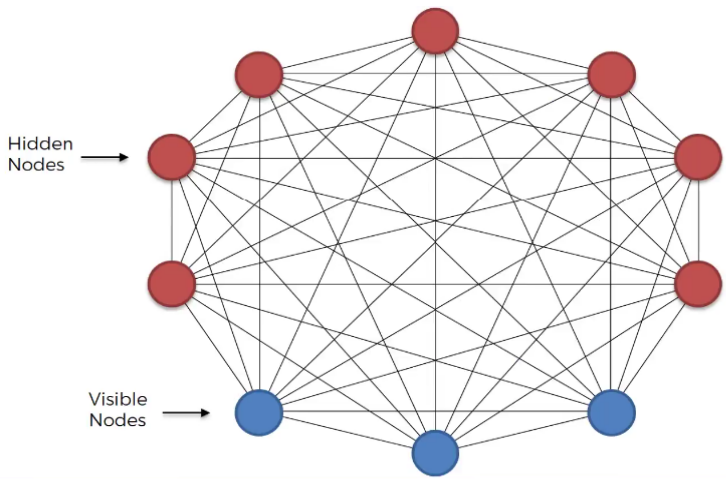

Fig. 23: Boltzmann Machine.

3.4.5 Restricted Boltzmann machine (RBM)

RBM was first presented to the public in 1986 by Smolensky. It consists of two levels, visible and hidden units, however there is no link between the visible and the secret layers. It is able to discover a probability distribution across a group of datasets to which it is exposed. Learning new features, collaborative filtering, reducing the number of dimensions, and classifying data are some of the uses of RBM [20].

3.4.6 Convolutional Neural network (CNN)

In CNN, each layer has a sensitive connection not just to the one below it but also to the other levels. Each neuron in the succeeding layer has a distinct purpose, such as being accountable for just a portion of the information coming in from the layer below it. Nowadays, remote sensing, computer vision, audio processing, and text processing are among CNN's most common applications (Fig. 24), [20], [31].

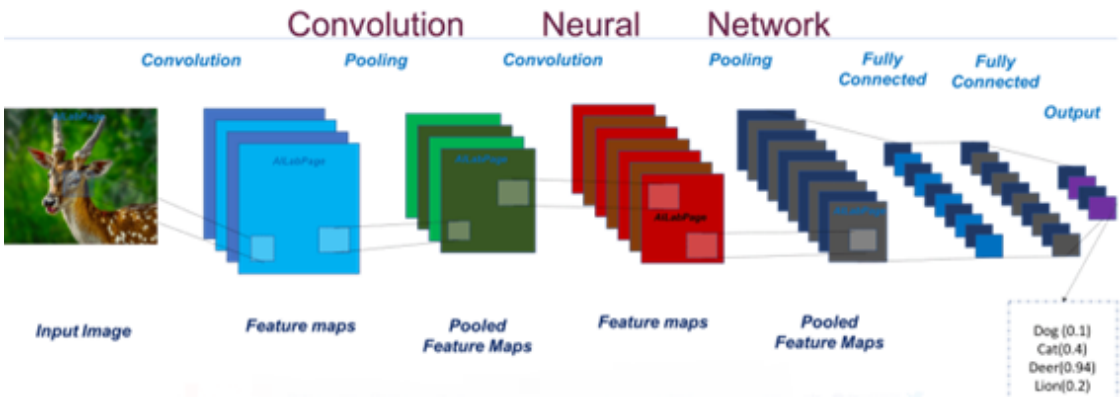

Fig. 24: Convolutional Neural Network.

3.4.7 Deep auto-encoder

The deep auto-encoder, like the others, contains numerous layers that are concealed from view. The difference between a basic auto-encoder and a deep auto-encoder is that the simple auto-encoder only has one hidden layer, but the deep auto-encoder contains a large number of hidden layers. Training a deep auto-encoder is often rather difficult since you have to train one hidden layer initially in order to reconstruct the structure of the input data. These input data are then used to train other hidden layers, and so on. Picture extraction, image creation recommendation systems, and sequence to sequence prediction are some of the applications that deep auto-encoder may be used for [20].

3.4.8 Gradient descent (GD)

GD is an optimization method that is frequently used in machine learning to determine coefficient functions. Its primary purpose is to minimize the overall cost function, and it is also used in the process of reducing the overall cost. In situations when it is not feasible to estimate the parameters analytically, GD is used in order to determine the parameters that are sought. The GD weight of the model is modified after each epoch that passes using it. Guided machine learning is accomplished with its help [20].

3.4.9 Stochastic gradient descent (SGD)

However, GD is used when the datasets are small, whereas SGD is typically used when the datasets are large, and SD becomes very costly if it is used for a large number of datasets. Both GD and SGD are optimization algorithms; however, GD is used when the datasets are small, whereas SGD is used when the datasets are large [20].

3.4.10 Summary

Unsupervised learning techniques have been studied for a shorter period of time than supervised learning methods because they are more challenging to execute.
CNN, which is a kind of algorithm implemented in deep learning, is one of the most extensively utilized medical procedures for analyzing cardiotocographic tracings.
This will be further discussed in the next chapter.

CHAPTER FOUR: UNSUPERVISED DEEP LEARNING METHODS IN CARDIOTOCOGRAPHY: LITERATURE REVIEW

4.1 Methods

This chapter compare some articles that use unsupervised learning as a method of analyzing CTG tracings for the purpose of detecting fetal hypoxia as early as possible.

Looking at various recently published papers, it is evident that until a few years ago, machine learning techniques were used for these types of studies (Fergus and Hussian's study in 2017 is a good example [32]), and then these studies led to an increase in the use of deep learning techniques as the need for improvements in this area has begun to be felt.

PubMed, Google Scholar, Elsevier, Springer, and the National Center for Biotechnology Information are the search engines that were utilized for the purpose of searching numerous research that focused on the application of deep learning algorithms for CTG traces.

Because locating articles that were as recent as possible was a special goal of the search, it only considered those that were published no more than four or five years ago and excluded all other results in order to make the search more up to date.

Articles that did not take into account the FHR signal or UCs were excluded from the study, as were all studies that used supervised learning types. However, articles that compared various types of learning were considered to also take into account the significant differences that exist between each type, which allowed for conclusions to be drawn.

In addition, preferences were communicated throughout the search process by making use of various keywords. These keywords included "unsupervised deep learning", "unlabelled data", "deep neural network", "cardiotocography", "fetal distress", "fetal hypoxia", "FHR", and "CTG tracings".

4.2 Results

The numerous studies that were taken into consideration may be found in this part. The study that was done was done utilizing the criteria described in the paragraph before this one. There is a total of seven of them, and they were all published within the past few years. The outcomes of all of them are not necessarily very positive (in fact, for some of them, there is space for improvement) but a significant amount of information that can be applied to upcoming innovations can be acquired from these publications.

4.2.1 Petrozziello's et al. study

"Deep Learning for Continuous Electronic Fetal Monitoring in Labor" by Petrozziello et al. [33] is a study conducted in 2018 that utilize raw information from Electronic Fetal Monitoring (EFM) to predict fetal discomfort.

The fact that computerized evaluation of EFM has not previously shown any advantage in randomized treatment trials is the novelty that this research aims to demonstrate. They offer data-driven computer techniques for EFM interpretation, such as deep learning, which enables automated evaluation based on massive clinical datasets.
They input the FHR and UC signals into a LSTM and a CNN network. They achieved a level of predicted accuracy that was 61% and 68%, respectively. It is important to point out that while their dataset included 35429 recordings, it also included 33959 healthy neonates and only 1470 compromised neonates, resulting in a very uneven collection of data. However, they demonstrate that CNN performs better than LSTM in predicting fetal impairment with false-positive rates that are either comparable to or lower. Additionally, they demonstrate that increasing the size of the training set improves the sensitivity and stability of CNN performance on the test set. When compared to other published feature extraction-based approaches, the performance of CNN shows only a slight improvement when evaluated on an external open-access database that is quite small.

4.2.2 Iraji's study

Iraji [34] in 2019 investigated additional soft computing techniques for fetal status prediction using cardiotocographic recordings in his study: "Prediction of fetal state

from the cardiotocogram recordings using neural network models". Neural networks and sparse stacking deep auto-encoders were employed in this respect. Iraji used a small dataset consisting of 2126 recordings that were sorted into three categories: 1655 normal, 295 suspicious, and 176 abnormal.

Iraji offers a mix of computer vision and soft computing methodologies in clinical decision-making by using training data. This demonstrates that Iraji is capable of improving both medical judgments and therapies. The fact that visual evaluation by clinicians may be both time consuming and erroneous served as the impetus for this particular research endeavor. As a result, the development of an intelligent computer system to evaluate the state of the fetus prior to the onset of labor in the mother is highly significant.

Within the scope of this research project, a number of different multilayer architectural topologies for a sub-adaptive neuro fuzzy inference system (ANFIS) using deep stacked sparse self-coders (DSSAE) have been suggested. The findings of the experiments reveal that DSSAEs are more accurate than other strategies that have been presented for predicting the status of the fetus.

On the overall dataset, DSSAE performed the best with a 96.7% accuracy rate, followed by ANFIS with a 95.3% accuracy rate.

4.2.3 Zhao's et al. study

Zhao et al. [35] during 2019 used FHR signals modified using the continuous transform to create pictures in their study: "DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network". Their dataset included the open-access database (CTU-UHB), which included 552 intrapartum FHR recordings and a sizeable proportion (about 20%) of scalp electrode recordings (SER). They offer an 8-layer deep convolutional neural network (Deep-CNN) that can automatically detect fetal acidemia in this body of work. The continuous wavelet transform (CWT), which provides a better way to observe and capture the hidden characteristic information of FHR signals in both the time and frequency domains, is used to obtain two-dimensional (2D) input images. This allows the information to be observed and captured more accurately. It is quite important to note that in contrast to traditional methods to machine learning, this line of work does not involve the performance of intricate feature

engineering, also known as the extraction and selection of features. In point of fact, the 2D CNN model is able to auto-learn relevant features from the input data with the condition of not losing informative features. This is the significant advantage that deep learning has over machine learning. In in fact, by using the CNN setup, they are able to attain superior classification performance when compared to other approaches. Their model has an AUC of 97.82% and an accuracy of 98.34%.

4.2.4 Rahmayanti's et al. study

"Comparison of machine learning algorithms to classify fetal health using cardiotocogram data" is very recent research published in 2022 by Rahmayanti et al. [36]. By using 21 parameters resulting from the measurement of FHR and UC, they compare learning algorithms for identifying fetal well-being. They claim high degrees of precision. The utilized dataset was collected from the University of California Irvine Machine Learning Repository, a public dataset. It includes of 2126 data obtained via the Sys Porto system from pregnant women in their third trimester. It is very interesting that this work uses the same database as this thesis work. Additionally, it compares seven algorithms (using three different scenarios) for predicting fetal health: Artificial Neural Network (ANN), LSTM, XG Boost (XGB), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Light GBM (LGBM), and Random Forest (RF). This work is very interesting because it uses the same database as this thesis work. The findings indicate that five of the seven algorithms have extremely excellent performance (accuracy ranging from 89 to 99%). XGB, SVM, KNN, LGBM, and RF are the names of these five algorithms. In addition, only one of the five algorithms, namely LGBM, performs well in all three circumstances on a constant basis. Moreover, in their investigation, the implementation of deep learning approaches did not provide outcomes that were superior to those of the ML strategy. The authors believe that performance can be enhanced by using a more representative database and refining the set of variables.

4.2.5 Spairani's et al. study

Spairani's paper [37] (2022), which was included in a recently published research by the Politecnico di Milano, is unquestionably worthy of consideration. The title of the study is "A deep learning mixed-data type approach for the classification of FHR signals".

The research at hand presents a brain architecture-based hybrid strategy for classifying healthy and sick fetuses based on a collection of quantitative data and images. Combining quantitative regressors expressing fetal health with attributes implicitly derived from many images (representations of FHR).

Consequently, there is a neural model with two linked branches, each of which consists of a multilayer perceptron (MLP) and a CNN. The neural brain architecture was trained using a large and balanced collection of clinical data (14000 CTG tracings, including 7000 healthy and 7000 diseased infants) obtained during ambulatory non-stress testing at the Azienda Ospedaliera Universitaria Federico II Hospital in Naples, Italy.

The results rely heavily on the number of cases, the database used, the attributes examined, and the database utilized, the characteristics considered, and the performance of the classifiers. In any case, following hyperparameter tweaking and training, the suggested neural network attained an overall accuracy of 80.1%, which is an encouraging result given that it was gained on a massive dataset.

4.2.6 Asfaw's et al. study

"Multimodal deep learning for predicting adverse birth outcomes based on early labour data", the research suggested by Asfaw et al. in 2022 [38], utilizes CTGs gathered from 51449 term babies to categorize those with and without severe birth damage based on the first 20 minutes of FHR recordings done in accordance with the standard of care at a big UK tertiary hospital over a 25-year period.

Three distinct 1D-CNN and LSTM-based architectures are compared: 1D-CNN, sequential 1D-CNN-LSTM, and parallel 1D-CNN-LSTM. Moreover, a 2D-CNN architecture is used to process 2D pictures. Parallel connections are made between the 2D-CNN and 1D-CN-LSTM in the proposed multimodal architecture. The partial area under the curve (PAUC) between 0-10% false-positive rate and the sensitivity at 95% specificity are used to assess the performance of the models. The findings reveal that the parallel 1D-CNN-LSTM design beats the sequential 1D-CNN-LSTM architecture and the multimodal architecture, reaching a PAUC of 20% and a sensitivity of 20% at 95% specificity. These findings are highly promising from a clinical standpoint, since they provide risk categorization from the beginning of monitoring and labor and early identification of as

many “at risk” people as feasible; nevertheless, implementation would need longer FHR traces and a larger data collection.

4.2.7 Feng’s et al. study

This paper by Feng et al. [39], somewhat less recent than the others but quite current (2018), is very interesting in that it proposes a comparison between supervised and unsupervised learning. The title of the paper is “Supervised and Unsupervised Learning of Fetal Heart Rate Tracings with Deep Gaussian Processes”. The performance of DGP (unsupervised) is compared to that of a supervised machine learning technique, the support vector machine (SVM), in identifying FHR traces (in the first case) and FHR traces mixed with uterine activity (UA), (in the second case). Utilized is a database including 552 CTG recordings obtained from the obstetrics department of the University Hospital of Brno, Czech Republic. The categorization is pH-based, and the unsupervised learning network is a five-layer DGP network. Experiments demonstrate that the DGP outperforms the SVM-based technique. When employing both FHR and UA capabilities, the DGP achieves greater performance than when using just FHR features. In fact, the sensitivity of DGP is greater than that of SVM in the analysis of FHR and UA signals (92% of DGP compared to 82% of SVM).

From this article, it can be determined that the most innovative path ahead, while not yet finalized, is to adopt deep learning algorithms based on unsupervised learning to eliminate the extremely variable and hence error-prone interpretation of CTG tracings by obstetricians.

4.3 Discussion

Reference	Author	Year	Database	Input	Algorithm Types	Results Obtained
33	Petrozziello et al.	2018	35429 CTG traces (33959 healty, 1470 compromised infants)	FHR, UC	LSTM, CNN	LSTM accuracy 61%, CNN accuracy 68%
34	Iraji	2019	2126 CTG signals (1655 normal, 295 suspect, 176 pathological)	FHR, UC	deep auto-encoders, ANFIS	deep auto-encoders accuracy 96,7%, ANFIS accuracy 95,3%
35	Zhao et al.	2019	552 intrapartum recordings	FHR, SER	CNN	AUC 97,82%, accuracy 98,34%
36	Rahmayanti et al.	2022	2126 CTG signals (1655 normal, 295 suspect, 176 pathological)	FHR, UC	ANN, LSTM, XGB, SVM, KNN, LGBM, RF	only LGBM obtained good results
37	Spairani et al.	2022	14000 CTG tracings (7000 healty, 7000 diseased infants)	FHR	MLP+CNN	accuracy 80,1%
38	Asfaw et al.	2022	51449 term babies	FHR	parallel 1D-CNN-LSTM, sequential 1D-CNN-LSTM	PAUC 20%, sensitivity 20% at 95% specificity
39	Feng et al.	2018	552 CTG recordings	FHR, FHR+UA	DGP, SVM	SVM (FHR) specificity 82%, sensitivity 73%; SVM (FHR+UA) specificity 82%, sensitivitity 82% DGP (FHR) specificity 91%, sensitivity 73%; DGP (FHR+UA) specificity 82%, sensitivitity 91%

Table 3: Comparison of studies based on deep learning and CTG tracings aimed at finding fetal hypoxia.

This kind of selection provides some insight into the background of CTG trace analysis as well as its development through time via the use of various forms of learning. In general, the published publications demonstrate that the accuracy and overall performance of AI algorithms for prenatal medicine are virtually inversely related to the number of cases: the best results are achieved with restricted and well chosen datasets. In fact, the vast quantity of data sets necessary to train neural networks is the primary drawback of deep learning approaches.

Consequently, selecting an insufficient sample size might result in an overestimation of the model's generalization capabilities.

Another argument that may be made is that it has lately been shown id that also hybrid models seem to perform good results. This is something that can be argued.

Table 3 was developed so in order to more effectively discuss and compare the different findings. The assessment and comparison of the seven studies that were taken into consideration are shown across seven columns in the table, which has the following organization: the reference number, the author, the year, the database that was utilized, the input, the kind of algorithm that was applied, and then lastly the results that were achieved.

It is important to first point out that all of the research are quite recent, given that the use of deep learning algorithms is relatively new. In this context, it is important to mention that the results achieved by using deep learning algorithms are superior than those acquired by using machine learning methods. This may be shown, for instance, in the research that Iraji and Feng conducted.

It is also clear that the FHR signal is used as the input for all studies with the exception of three, one of which additionally employs SER in addition to the FHR signal.

Another thing that has been noticed is that the database that was used has a significant impact on the results that were obtained. The development of deep learning has made it possible to use larger databases, particularly when it comes to unsupervised learning; however, there is still room for improvement in terms of how the system handles large amounts of data, as this can have a negative impact on the outcomes. It has been shown that databases containing a limited amount of data are capable of producing results with a greater degree of precision.

CNN is the most used deep learning algorithm, and it works well in conjunction with DGP to provide satisfactory results.

The last study represented in the table is a direct comparison of unsupervised (DP) learning to supervised (ML) learning, and the results demonstrate that unsupervised learning is preferable since it achieves higher levels of sensitivity and specificity.

Nevertheless, each of these described experiments obtains great findings that are easily replicable and leaves a lot of space for improvement for further advances.

CHAPTER FIVE: UNSUPERVISED DEEP NEURAL NETWORK FOR FETAL CLASSIFICATION

5.1 Database

The dataset that was utilized was acquired from the Faculty of Medicine at the University of Porto, in Portugal. This dataset is available to the public. It includes information from 2126 pregnant women who were in the third trimester of their pregnancies when the survey was taken. These 2126 CTGs went through an automated processing system, and the diagnostic features of each individual CTG were evaluated. SisPorto 2.0 (Speculum, Lisbon, Portugal), a tool that performs automated analysis of CTG findings, is the one responsible for the processing of fetal CTGS automatically [40]. Additionally, the CTGs were classified by three expert obstetricians, and each was given a designation that reflected the consensus of their categorization. The categorization was carried out with consideration given to both the morphological pattern (A, B, C, D…) and the fetal condition. Therefore, the dataset is suitable for use in studies involving 10 classes as well as 3 classes. In this study, it was used for 2-class studies, and as a result, the categorization of cases as N=normal, and P=pathological was taken into consideration.
Table 4 presents the 21 characteristics that are included in this dataset and are utilized for performing FHR and UC measurements on CTG.

ATTRIBUTES	
LBE	baseline value (medical expert)
LB	baseline value (SisPorto)
AC	accelerations (SisPorto)
FM	foetal movement (SisPorto)
UC	uterine contractions (SisPorto)
ASTV	percentage of time with abnormal short term variability (SisPorto)
mSTV	mean value of short term variability (SisPorto)
ALTV	percentage of time with abnormal long term variability (SisPorto)
mLTV	mean value of long term variability (SisPorto)
DL	light decelerations
DS	severe decelerations
DP	prologued decelerations
DR	repetitive decelerations
Width	histogram width
Min	low freq. of the histogram
Max	high freq. of the histogram
Nmax	number of histogram peaks
Nzeros	number of histogram zeros
Mode	histogram mode
Mean	histogram mean
Median	histogram median
Variance	histogram variance
Tendency	histogram tendency: -1=le60symmetricric; 0=symmetric; 1=right assymetric

Table 4: features of the dataset utilized.

5.2 Methods

The following is the approach that was followed in order to arrive at the construction of an unsupervised deep neural network that functions automatically on two different levels of categorization. The dataset was analyzed and separated from the raw data. To do this, the features collected were studied and the way they were extracted was investigated in depth. In the course of the research, only the two-level categorization was taken into consideration; the ten-level classification was excluded entirely. The N, and P classifications make up the foundation of the categorization system (N for normal, and P for pathological). It was determined, on the basis of the classifications, how diagnostically powerful each characteristic was. Specifically, an assessment was made to see whether the features had normal distribution, using the Lilliefors test. From this test, it emerged that the data did not have normal distribution; at this point, the median and interquartile range (IQR) of each feature in each of the two classes were calculated, as can be seen in Table 5. In addition, the Wilcoxon rank sum test was used in order to analyze the statistical significance of the comparisons made between P vs N.

At this point the dataset was divided into two datasets, one for training and one for testing. The division was done keeping in mind the classes to which they belonged, so to the training dataset went 70% of each of the two classes. The remaining 30% made up the test dataset. The distribution of classes was thus kept intact. After that came the development of the Neural Net Clustering (NNC) toolbox that is available on the Matlab machine learning platform. NNC was used for the purposes of generating and training a network of self-organizing maps (SOM). SOMs generated in this study are a form of neural network structure in which output neurons are arranged in 2D grids and all input neurons are linked to all output neurons. They are taught to construct a representation of the training samples in a 2D space while retaining the topological features of the input space using unsupervised learning. The ultimate goal of the implementation of SOMs is proposing two-level clustering. To be more specific, the training dataset was imported, the 21 inputs, the attributes, were chosen, and the number of neurons was set to ten. Next, Batch weight training was initiated, which is a method of training in which the weights and biases are updated only after all of the inputs have been presented. The

data were analyzed using mean squared error (mse). Fig. 25 illustrates the finished network that was established.

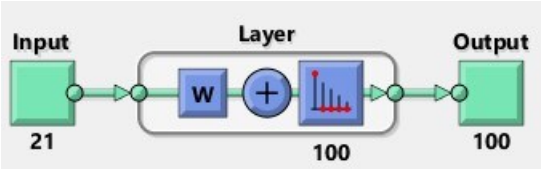

Fig. 25: Self-Organizing Map.

The network after being trained on the training database was tested on the test database.

Then, a vector quantization approach was used to partition the gathered observations into k clusters, with k being equal to 2 in this study. In this manner, each observation is assigned to the group that has the median value that is the closest (cluster centroid). Finally, a confusion matrix was used to analyze the number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) in order to calculate the accuracy of the results processed by the network with the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

5.3 Results

Observations may be drawn from the statistical analysis performed prior to the neural network building procedure.

First, Table 5 summarizes the outcomes of both tests conducted on the 21 accessible characteristics' data. In fact, we see that the first column includes all features, whereas the other two columns provide the median ± IQR for both classes (N and P), respectively. Additionally, the Wilcoxon rank sum test result is presented in the last column, and a p-value less than 0,05 is indicated by an asterisk next to any number that falls into this group. A p-value less than 0,05 shows that the statistical significance of the test on those features.

FEATURES	N	P
LB	132±13	132±6
AC	0,003±0,0067	0*
FM	0±0,02	0,0009±0,0022*
UC	0,0049±0,004	0,0033±0,0064*
ASTV	0±0,0035	0,0023±0,0067*
mSTV	0	0*
ALTV	0	0,0011±0,0023*
mLTV	41±27	65±12,5*
DL	1,3±0,975	1,7±2,2
DS	0±4	0±65,5*
DP	8±6,675	3,25±5,8*
DR	71±58	92,5±95
Width	90±47	66±68*
Min	163±23	158,5±32,5*
Max	4±4	4±6
Nmax	0	0±0,5
Nzeros	138±18	122,5±44*
Mode	135±19	106,5±38,5*
Mean	138±18	116±29*
Median	9±21	36,5±79*
Variance	0±1	0±1*

Table 5: Statistical analysis on all the features; the median ± IQR is indicated for each class; *characteristics with a p-value≤0,05 are indicated.

Fig. 26 represents the SOM sample hits of the training and test, respectively. In this picture, it can be seen the neural network training taking place during the training phase (on the left), as well as the neural network evaluating its abilities during the testing phase (on the right). The blue spots denote a particular form of clustering, and the number that is written inside of each hexagon denotes the number of neurons that are connected to that particular cell. Because the training database has more information, the number of associations that occur in the test is less than the number that occur in the training database. While there are 1282 signals in the training database, there are only 549 signals in the test database.

After this point, the data will be utilized for classification on two levels in order to arrive at a BXC matrix (where B stands for babies and C stands for classes), which will then be examined using the confusion matrix.

Fig. 27 displays the outcomes of the training and testing procedures that were conducted using the confusion matrix. The real values may be found along the y-axis, while the values that were predicted can be found along the x-axis. From each confusion matrix it was computed the level of accuracy: for the training it is around 74%, for the test the accuracy is equal to 82%.

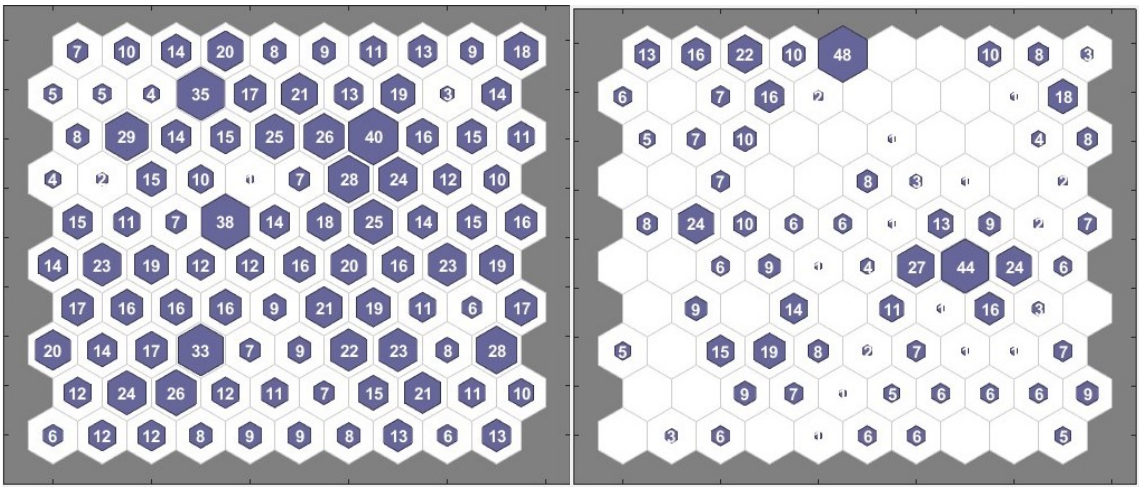

Fig. 26: SOM sample hits of the training (on the left) and test (on the right).

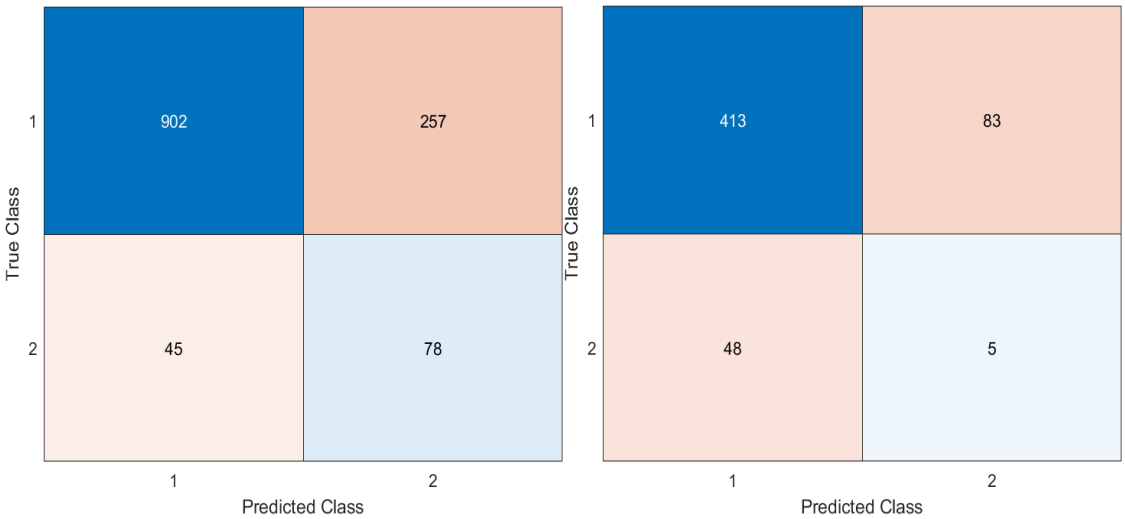

Fig. 27: confusion matrix of the training (on the left) and test (on the right).

5.4 Discussion

The IQR index is used to quantify the degree to which the values of the characteristics diverge from the sample median. Because of this, the value's magnitude will determine the amount of variation that exists from the median. In terms of the IQR, Table 5 reveals that the majority of the characteristics have a very low IQR, which, in a few cases, comes close to 10, but is still, by and large, rather low. This is the case despite the fact that some of the cases border on 10. Then there are certain characteristics that have a rather high IQR value (width, min, max mode, median and variance). It's interesting to note that these characteristics have a high IQR value for both N and P. The same is true for the values that have a very low IQR index; in point of fact, it can be observed that the values are rather comparable to one another and differ only little between the two groups. This indicates that the same behavior was seen in two separate settings.

The Wilcoxon test is the second one that is carried out, and its results are shown in the table. This test is performed so that we may establish whether or not the parameters have sufficient statistical significance. The values in the P column that are shown here are denoted with an asterisk. It is clear from this table that all of the values, with the exception of these four (LB, width, nmax, and nzeros), have a p-value that is lower than 0.05, which indicates that they are statistically significant.

As a consequence of these preliminary and statistical studies, positive findings have been achieved, which improves the choice of this database.

Looking at the neural network, on the other hand, it can be observed the manner of training that was carried out on the network (Fig. 27), as well as the testing phase that followed. When the network was being trained, it conducted clustering that seems to have been more than acceptable. By examining the picture in great detail, one can see that blobs, one for each class, were generated throughout this process. This indicates the capacity of the net to create clustering by associating the inputs received. This is also mirrored in the period of testing, and in fact, it is much more obvious there.

Finally, the findings that were applied in the confusion matrix present extremely encouraging news in that an accuracy of 74% was reached in the training phase and 82%

was gained in the testing phase. These findings are outstanding since they are superior to the results that are often achieved in tests that are quite comparable to this one.

It is clear from the confusion matrix that a relatively high number of TNs are acquired throughout testing and training; in addition, in testing phase there is a much smaller number of FP during the phase of training.

Based on these encouraging findings, the database seems to be rather strong, both in terms of its size and the class homogeneity it offers.

It is demonstrated that the classes can be distinguished from the network with minimal effort.

CONCLUSIONS

In general, this study recommends the use of intelligent methodologies in order to enhance the quality of care and diagnosis that is offered to medical professionals on the state of the fetus after birth. The purpose of this project is to enhance the accuracy of diagnostic procedures for fetal hypoxia state, with the end objective of eliminating mistakes brought on by the expertise of medical professionals and making diagnosis as quick and efficient as is humanly feasible.

In point of fact, the urgent need for a new route that may lessen the strain on the shoulders of medical workers and at the same time enhance the quality of life of the baby and his or her family is the impetus behind this line of study. It is possible that the adoption of diagnostic approaches that are based on deep learning might lessen the need for more specialist examinations. These systems can learn from tracings of previously diagnosed babies in order to identify new instances. These methos, which is based on deep learning, might also be used to educate non-specialist doctors on how to enhance their decision-making processes.

The fact that unsupervised learning was the type of learning that was utilized to undertake classification work is the most notable feature of the methodology that was used to this kind of work. That makes the findings even more mind-blowing: in point of fact, it is very novel that the network trained for this research was able to categorize on two levels of fetal health despite the fact that there was no correcting bias present. The accomplishment of 82% accuracy during testing is not in any way something that should be undervalued; in fact, it provides the groundwork for more research to be conducted in this sector and using the technique described above.

In fact, accuracy was the method that was used in order to produce an assessment of the performance of the classifier.

In addition, within the scope of this investigation, a variety of different categorization strategies and databases were investigated (by analyzing recent articles). One point that can be made is that while a database with the size that was indicated before has the potential to be useful in order to acquire results that are on the verge of being accurate, there is still space for improvement. One additional item that has come to light is the

fact that the approaches of deep learning are unquestionably more successful than the ones that were used in the past. On the other hand, the unsupervised type of learning represents the vanguard, and this study confirms that if it is deepened, results can be obtained that can lead to the daily use in the medical field of this methodology to detect fetal hypoxia. In spite of the fact that making a significant difference between normal and pathological tracings seems to be by far the most decisive and effective manner, one enhancement that can undoubtedly be made is the implementation of the third class, which is that of suspicious instances.

BIBLIOGRAPHY

[1] Talaulikar V, Arulkumaran S. Maternal, perinatal and long-term outcomes after assisted reproductive techniques (ART): implications for clinical practice. Eur J Obstet Gynecol Reprod Biol (2013);170:13-9.

[2] Bobrow C, Soothill P. Causes and consequences of fetal acidosis. Archives of Disease in Childhood-Fetal and Neonatal Edition (1999);80:F246-F249.

[3] Hasan M, Reaz M, Ibrahimy M, Hussain M, Uddin J. Detection and processing techniques of FECG signal for fetal monitoring. Biological procedures online (2009);11:263.

[4] Rurak D. Fetal oxygenation, carbon dioxide homeostasis, and acid-base balance: New York: Oxford University Press (1994).

[5] Singer D. Neonatal tolerance to hypoxia: a comparative-physiological approach. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology (1999);123:221-234.

[6] Godfrey K, Barker D. Fetal nutrition and adult disease. The American journal of clinical nutrition (2000);71:1344s-1352s.

[7] Pattar R, Raj A, Yelamali B. Incidence of multiorgan dysfunction in perinatal asphyxia. International Journal of Contemporary Pediatrics (2017);2:428-432.

[8] Rounaq Abdalrahman A. Classification of foetal distress and hypoxia using machine learning. Liverpool John Moores University (2018);23-48.

[9] Itskovitz J, LaGamma E, Rudolph A. The effect of reducing umbilical blood flow on fetal oxygenation. Am J Obstet Gynecol (1983);145:813-18.

[10] Parer J, Livingston E. What is fetal distress?. International Journal of Gynecology & Obstetrics (1991); 34:389-389.

[11] Manning F, Platt L, Sipos L. Antepartum fetal evaluation: development of a fetal biophysical profile. American journal of obstetrics and gynecology (1980);136:787-795.

[12] Thacker S, Berkelman R. Assessing the diagnostic accuracy and efficacy of selected antepartum fetal surveillance techniques. Obstetrical & gynecological survey (1986);41:121-141.

[13] Marcdante K, Kliegman R. Nelson Essentials of Pediatrics. Elsevier Health Sciences (2014).

[14] Butt T, Farooqui R, Khan M. Risk factors for hypoxic ischemic encephalopathy in children. J Coll Physicians Surg Pak (2008);18:428-432.

[15] Chudacek V, Spilka J, Janku P, Koucky M, Lhotska L, Huptych M. Automatic evaluation of intrapartum fetal heart rate recordings: a comprehensive analysis of useful features. Physiol Meas (2011);32:1347-60.

[16] Shahad N. Al-Yousif M. Mohd A. Cardiotocography Trace Pattern Evaluation Using MATLAB Program. Department of Electrical Electronics and System Engineering, University Kebangsaan MalaysiA (2009);153-157.

[17] Alfirevic Z, Devane D, Gyte G. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour (Review), The Cochrane Library 4, (2007).

[18] Amir Sweha M. Interpretation of the Electronic Fetal Heart Rate During Labor. Am Fam Physician (1999); 59(9):2487-2500.

[19] https://docs.microsoft.com/en-us/azure/machine-learning/concept-deep-learning-vs-machine-learning

[20] Saadat N, Shuaib M. Advancements Advancements in Deep Learning Theory and Applications: Perspective in 2020 and beyond. BKCI Book Citation Index in Web of Science™ Core Collection (2020);3-7.

[21] Ayodele T. Types of machine learning algorithms. New advances in machine learning. InTech (2010).

[22] Mohri M, Rostamizadeh A, Talwalkar A. Foundations of machine learning: MIT press (2012).
[23] Zhu X, Goldberg A. Introduction to semi-supervised learning. Synthesis lectures on artificial intelligence and machine learning (2009); 3: 1-130.
[24] Chapelle O, Scholkopf B, Zien A. Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. IEEE Transactions on Neural Networks (2009);20:542-542.
[25] Sutton R, Barto A. Reinforcement learning: An introduction vol. 1: MIT press Cambridge (1998).
[26] Sutton R, Barto A. Introduction to reinforcement learning vol. 135: MIT Press Cambridge (1998).
[27] https://www.ai4business.it/intelligenza-artificiale/deep-learning/deep-learning-cose/
[28] Zhong S, Hu J, Fan X, Yu X, Zhang H. A deep neural network combined with molecular fingerprints (DNN-MF) to develop predictive models for hydroxyl radical rate constants of water contaminants. Journal of Hazardous Materials (2020);383(5).
[29] Zhang Z, Zheng L, Wang M. An exponential-enhancedtype varying-parameter RNN for solving time-varying matrix inversion. Neurocomputing (2019);338:126-138.
[30] Kong J, Huang J, Yu H, Deng H, Gong J, Chen H. RNN-based default logic for route planning in urban environments. Neurocomputing (2019);338:307-320.
[31] Konstantinidis D, Argyriou V, Stathaki T. A modular CNN-based building detector for remote sensing images. Computer Networks (2020);138: 107034.
[32] Fergus P, Hussain A, Al-Jumeily D, Huang D, Bouguila N. Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. Biomed. Eng. OnLine (2017);16: 89.
[33] Petrozziello A, Jordanov I, Papageorghiou T, Redman W; Georgieva A. Deep learning for continuous electronic fetal monitoring in labor. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (2018); 5866–5869.
[34] Iraji M. Prediction of fetal state from the cardiotocogram recordings using neural network models. Artif. Intell. Med (2019);96:33–44.
[35] Zhao Z, Deng Y, Zhang Y, Zhang X, Shao L. DeepFHR: intelligent prediction of fetal acidemia using fetal heart rate signals based on convolutional neural network. BMC Med. Inf. Decis. Mak (2019);19:286.
[36] Rahmayanti N, Pradani H, Pahlawan M, Vinarti R. Comparison of machine learning algorithms to classify fetal health using cardiotocogram data. Procedia Comput. Sci (2022);197:162–171.
[37] Spairani E, Daniele B, Signorini MG, Magenes G. A deep learning mixed-data type approach for the classification of FHR signals. Frontiers in Bioengineering and Biotechnology (2022);2-4.
[38] Asfaw D, Jordanov I, Impey L, et al. Multimodal deep learning for predicting adverse birth outcomes based on early labour data (2022);2-18.
[39] Guanchao F, Quirk J, Djuri´c P. Supervised and Unsupervised Learning of Fetal Heart Rate Tracings with Deep Gaussian Processes. NEUREL Symposium on Neural Networks and Applications (2018);1-6.
[40] Ayres de Campos et al. (2000) SisPorto 2.0 A Program for Automated Analysis of Cardiotocograms. J Matern Fetal Med 5:311-318.