Towards an AI driven early detection of brain injuries in neonates through non-contact audio and video recording

Janet Pigueiras-del-Real[1], Lionel C. Gontard[1,4], Simón P. Lubián-López[2], Isabel Benavente-Fernández[2] and Angel Ruiz-Zafra[3]

[1]*Applied Optics and Magnetism Research Group, University of Cádiz, 11510 Puerto Real, Spain*

[2]*Department of Paediatrics and with Biomedical Research and Innovation Institute of Cadiz (INiBICA) Research Unit, Puerta del Mar University, Cadiz, Spain*

[3]*Department of Software Engineering, University of Granada, Periodista Daniel Saucedo Aranda s/n, 18071, Granada*

[4]*IMEYMAT, University of Cádiz, 11510 Puerto Real, Spain*

Abstract

An early detection of brain injuries in preterm infants' development fosters early therapies and treatments that could significantly improve the health of babies. Recent research confirm that the use of audio and video as non-contact data sources could enable the diagnosis of a possible brain damage of a neonate through the use of AI, but advances in this area are still very much in its infancy. This paper introduces an approach for the design and validation of a non-contact monitoring system to be used in a Neonatal Intensive Care Unit (NICU) that would help to the early detection of neonates affected by brain injury. The research focuses on the identification of neurological injury markers through the development of AI-based techniques based on video and audio data, exploiting the different features related to the movements, crying and sounds, and vital signs data of healthy neonates and of those affected by a brain injury. The paper presents the methodology and focuses on the first stage (System deployment) where it is described a software platform designed to collect, record and label data from different video and audio sources in a NICU, including the physiological parameters of the neonates.

Keywords

methodology, monitoring device, preterm infants, audio and video technologies, artificial intelligence, video and data gathering

1. Introduction

Early diagnosis of problems that can lead to neurodevelopmental disorders in preterm neonates are considered one of the main concerns of the medical community [1]. These patients have some vital functions immature and they need special care in Neonatal Intensive Care Units (NICUs) where their physiological parameters (heart rate, respiration rate and oxygen saturation) are constantly monitored with wired sensors attached to the skin of the baby. These sensors are connected to a monitor screen, which is reviewed by medical staff. Although this technique

DETERMINED 2022: Neurodevelopmental Impairments in Preterm Children — Computational Advancements, August 26, 2022, Ljubljana, Slovenia

janet.pigueiras@uca.es (J. Pigueiras-del-Real); lionel.cervera@uca.es (L. C. Gontard); isabel.benavente@uca.es (I. Benavente-Fernández); angelr@ugr.es (A. Ruiz-Zafra)

is appropriate, use of non-invasive techniques are emerging in recent years as a solid and viable alternative [2, 3].

Clinical monitoring of preterm infants including by direct observation of motor activity, facial expression, skin color, or cry can also be used to detect potential problems in the neurodevelopment of the baby. However, not all newborns can get benefit from this clinical follow-up as it requires medical staff with especial training that it is not always available and if present they have to share their time among many babies in a NICU [4].

The collection and analysis of audio and video of the neonates has proven to be a feasible solution with advantages for their monitoring, since it offers the gathering of clinical data without the need to use invasive methods (sensors glued to the skin) that may lead to discomfort and stress periods for the preterm. These techniques are currently being widely used in other biomedical applications [4].

In this paper we present a draft of an approach for the early detection of brain injuries in neonates using audio and video and supported by different computing technologies and paradigms: artificial intelligence (AI), internet of things (IoT), edge computing and computer vision (CV). The goal of the proposal is the use of non-invasive technologies to gather information related to the neonate through audio and video recordings, and process this data sources along with physiological information through the implementation of novel AI and CV-based algorithms/models to detect related symptoms of brain injuries in early stages of the baby's development.

The paper introduces the approach showing the different 5 stages that compose it, and focuses in the first two, which described *1)* an edge computing-based device developed and deployed to gather audio/video information in NICUs and *2)* the description of the *Neonate Recording Platform*, or NRP, A recording software platform used to collect data from different sources: video cameras, microphones and also from the monitors that are used in NICU for visualizing the physiological parameters of the babies.

The paper has 4 sections and it is organized as follows. In section 2 we present a detailed literature review. Section 3 illustrates the proposed methodology, with the different stages of this approach and the Section 4 provides conclusion of the study.

2. Related work

The recording of video and audio enables a non-invasive way to collect relevant medical information of the patients and it is being applied in many biomedical domains [4]. Next, we review recent developments for the non-contact monitorization of preterm neonates.

Video

In the case of NICUs, the research presented in [5] measures with video physiological variables and automatically detect bradycardia in infants. Other authors [6, 7] have developed and improved algorithms to control oxygen saturation, showing that it is safe and effective for carrying out measurements of vital parameters in preterm infants with assisted mechanical ventilation. Other studies have proposed the use of computer vision for the identification of sleep stages, combining information from eye movements, body movements, facial expressions, sound made by babies and breathing patterns [8, 9].

A trendy line of research in this area is the automatic analysis of video using artificial intelligence techniques with the aim to provide an early diagnosis of neurological disorders [10, 11, 12]. *AI-based* models are used to detect the baby's pose and group them (set of sequences) to check if it is a normal behavior or abnormal. For instance, it has been applied to detect signs of cerebral palsy in babies [13, 14]. Much research focuses on the investigation in babies with a gestational age of 36 weeks, in order to correlate the amount of movement of the body with pain [12]. In the same category, in the project presented in [15] the authors propose to analyse spontaneous movements in order to diagnose neurodevelopmental disorders. Although this area is of great interest also for neonates, most of the research related to study the sequences of movements of the baby is carried out with a population of children who already walk.

Audio

Regarding to the use of audio as data source, the investigations in paediatrics mostly relied on cry analysis. In the 2000s, the analysis of signals began to be automated thanks to the *AI* techniques [4]. There are research that addressed the classification of cry signal through the use of *AI* algorithms in order to determine when babies are hungry, sleepy, need attention, are uncomfortable or need a diaper change [16, 17]. Analysis of cries was also developed in other contexts, for instead, [18] shows that deep learning systems are a powerful machines that can be used for distinguishing between healthy and pathological infant cry records. The authors of these works state that the crying of the baby is a field that has not yet been widely explored since it is not a language that can be easily understood, despite the fact that it is the main means of communication for this population.

In [16, 18] the authors made used of the short-time Fourier transform (STFT) to analyze audio signals. They also apply techniques originally designed and used in automatic speech recognition to detect and recognize the features of the baby's cry, and compression sampling to analyze and classify these signals. In addition, other research apply tools that were developed for analysing the crying signal. We can mentioned the study of [17, 19]. The authors apply the BioVoice software tool, developed specifically for the acoustic analysis of a newborn audio signal, in order to address their work in the crying field.

Furthermore, another topic in the use of audio for neonates is the measurement and analysis of environmental noise in the NICU. Probably [20] is the current state of art in this topic, where the authors develop the automatic detection of acoustic alarms in a noisy environment by applying filtering in the frequency space.

It is important to notice that joint audio and video processing has not been widely addressed so far. Concerning to our knowledge, only one study that integrates audio and video processing was published in [21] with Digi-New B project, where non-invasive strategies are proposed for the early diagnosis of neonatal sepsis.

In this paper we present an approach that joints the use of audio and video for a non-invansive detection of brain injury in neonates in early stages. The rest of the paper presents the different stages of the approach and focus in the technology and main contributions developed to collect audio and video data.

3. The proposed approach

This section introduces the draft of the proposed approach, which consist in a methodology composed by five stages and different AI-based techniques, developed computer vision algorithms, etc. The five stages of the methodology are depicted in Figure 1 and each stage is described shortly thereafter.

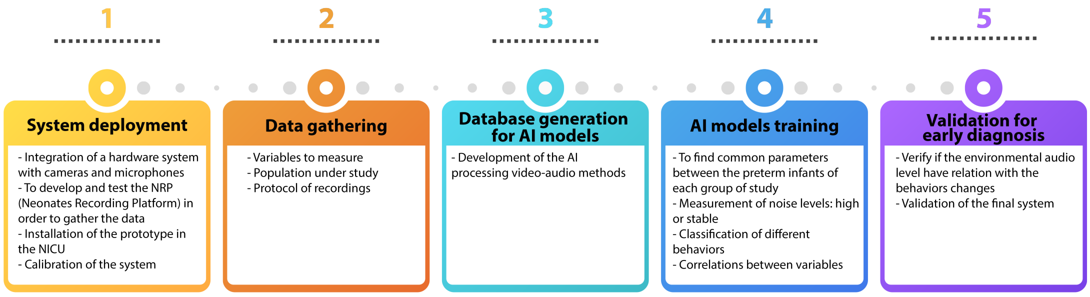

Figure 1: Proposed approach

Stage 1: *System deployment*

As first stage of the methodology, we need to set up the hardware-software system that must be used to record video as well as audio. Several commercials recording devices are available, however, we define a set of requirements that must accomplished by the devices, according to our expected outcomes once the approach completely implemented. The device required in our approach should meet, at least, the following features:

- *High quality recording video device.* In order to process in the best possible way the different frames captured of the neonate, we envise the use of a high quality recording camera.
- *Depth sensor supports in video recording.* The video captured will be used by AI-based techniques, so it is also desired to have a camera that also records or calculate depth in the recording raw video.
- *Physiological parameters collection.* So far, this approach does not address the automatic detection of some physiological parameters of the neonate such as heart rate, respiration rate, oxygen saturation, etc. In this way, our recording device will also have an additional camera to record the device where these neonate's parameters are shown, that is, the medical device located near to the neonate where are the wired sensors are connected. The software of these medical devices is not open source and the medical staff do not have the possibility of exporting the physiological parameters values.
- *Multiple audio recording.* At least, our approach must cover the recording audio about the sounds of the neonate but also the environment. However, could be interesting support an unknown number of recording audio devices in order to add as many data sources of audio as needed.

- *Data connectivity.* Although the recording of audio and video can be stored anywhere locally, could be interested have a device that can store the recordings but also send all the data (or part of it) to an external data warehouse.
- *Additional external storage.* Most of devices currently has an embedded internal storage or a free slot to insert a microSD card. Sometimes, especially if we record many hours, this storage capacity is not enough, so it is desired also that the devices has an extension to add additional storage devices like hard drives disks.
- *Comfortably and handy devices.* This recording device will be used probably by a medical staff or a research, that is, by a single user. So, the device must be easy to set up and easy to use.

With this restrictions about our needs in our approach, and after a concise search, we have not found an available commercial device that meets these requirements.

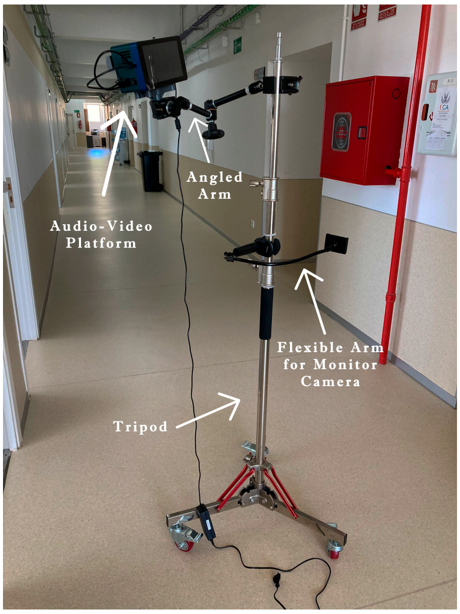

Figure 2: Prototype Support Equipment

In this first stage we have design, assembled and set up a hardware system with commercial components for the gathering of audio and video data from neonates, through the use of cameras and microphones. The system will record color videos with a camera focused on the baby, for gathering the body movements and facial expressions [22]. There will be a second camera that will be focused on recording the screen of the monitors that show the physiological parameters (heart rate, breathing, oxygen saturation) of the baby and that are placed next to the incubators. In addition the system will have 3 microphones, one omnidirectional to record environmental noise, and one directional to gather sounds made by the baby, and one additional from the physiological camera.

We have developed the system prototype (see Figure 2), that is composed of an OAK-D-CM4 (color camera), a Raspberry Pi where the operating system runs, an external hard drive, the camera to register physiological parameters (web camera) and, in the first term, three microphones (aforementioned). One of the main features of the OAK is that it runs any AI model, even custom architecture/built ones. We are going to need this characteristic for future development of the recording platform, to offer a diagnosis on the possible presence of neurological problems signs in a preterm infants (Stage 4 of this approach).

In addition, for the right placement of the recording prototype, we have to place it where does not disturb the work of the clinicians, and easy-to-handle for them. The 3D housing is held near the incubator with a tripod and an angled arm. The second camera is also attached to the tripod with a second flexible arm. The housing includes a touch screen for the execution of the NRP.

This prototype is fully hardware, so the software is still required. Although the prototype could be used in many different areas, this first version has been developed aligned to the goals of this approach. In this way, a customized software to use the prototype as well as satisfy the requirements is required.

We have developed a customized software called *Neonates Recording Platform (NRP)*. The main view of NRP is illustrated in Figure 3.

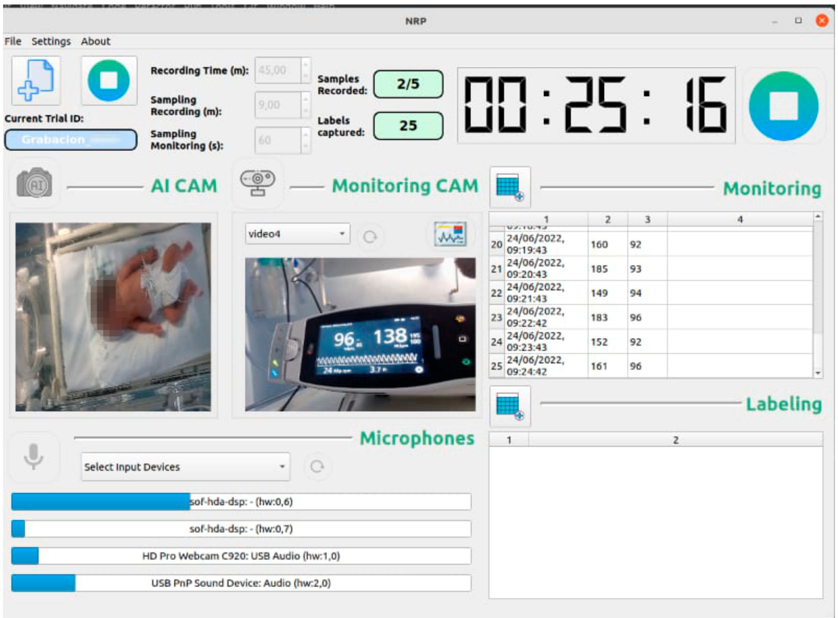

Figure 3: Neonate Recording Platform

The NRP (see Figure 3) supports the following different functionalities:

1. the recording of audio signals of different inputs/microphones. As many microphones as required can be added.
2. permanent recording of the neonate inside the incubator and the NICU (video).
3. video recording of the monitor that displays the neonate's physiological parameter.
4. automatic extraction and parsing of physiological parameters besides the entire video of the physiological parameters, a labeling system that automatically detects (using OCR

+ Machine learning neural network) from the video source *3)* and parse the neonate's parameters.

5. The software will enable healthcare professionals enabling to the addition of adding timestamps and comments
6. a trial-oriented structure to enable the exportation of the data in interoperable formats (XML, CSV, JSON) and enabling the exchange of trials between healthcare professionals
7. in order to ensure reliability, the recording interface supports the sampling of video and audio, recording content according to slots defined by end-user.

Stage 2: *Data gathering*

First of all, we need to calibrate the previous developed system in a real environment. For this purpose we are going to install the prototype in the NICU of the Puerta del Mar University Hospital located in Cadiz. The clinicians are going to choose one preterm infant and place him in an incubator with the same environmental conditions of the rest of the future recordings.

Secondly, we aim to be able to distinguish a healthy baby and a baby with brain injury. In order to achieve this goal, we need to analyse the different features of the behaviors of this 2 groups (crying, body movement and facial expressions), as well as the possible external triggers of the state changes. For this purpose, we propose to measure the following variables showed in the Table 1, that we have classified as external or internal variables depending on whether or not they are the baby's own.:

Table 1. Variables

Variable/Source	**Audio**	**Video**	**Other**
Internal	Crying	Movements	Heart frequency, Oxygen saturation, breathing frequency, clinical data
External	Environmental Audio	Presence/Absence of the baby	Comments and Labeling

The videos and audios will be recorded with the platform developed in the aforementioned Stage 1. At 36 weeks of post-menstrual age, the baby will be video-recorded for a period of 6 hours, 4hours before, and 2hours after the baby feeding. Around feeding time the baby is usually awake and we can gather his/her movements and crying. In addition, during this time controls are not usually carried out and the baby is not usually moved, so we can gather their movements and spontaneous crying without an external agent stimulating them.

The population under study will be made up of 2 groups of preterm infants, one of healthy babies and the other of babies with some neurological injury, who are admitted to the NICU of the University Hospital Puerta del Mar (Cadiz, Spain).

In order not to affect the possible movements of the baby, we propose to follow the protocol used for the authors [23, 24, 25] based on the Prechtl method [26] where it is proposed to carry out the recordings following the aspects:

- The child must remain in a supine position in the incubator
- Cosy incubator at a neutral temperature
- Free to move their body and all limbs, including fingers and toes
- Only wearing a diaper, no blankets or clothes
- Free disturbing environment

Stage 3: *Database generation for AI models*

Once all the data has been gathered, the pre-processing methods will be applied to provide the database for the training and later validation stage of the *AI* models.

The databases will save the different features that allow us to find the final correlation between variables in order to identify when a preterm present behaviors/state related to a possible brain injury. We could mention the following features:

- *Audio*: cry length, values of fundamental frequency (F0) and the first three resonant frequencies of the vocal tract (F1, F2 and F3), the decibels of environmental noise.
- *Movements/Motor activity*: value (angle and speed between adjacent joints) of the coordinates of 12 joints (left and right) shoulders, elbows, wrists, hips, knees and ankles.
- *Vital signs*: oxygen saturation value, breathing rate and heart rate.

Stage 4: *AI models training*

Through *AI* models, we aim to detect common data for each study groups that allow us to find behavioral features to differentiate a healthy baby from a baby with a possible neurological injury. In addition, classification methods will be applied in order to identify noise levels, high or stable.

On the other hand, through the application of the suitable *AI* model and using the data gathered previously, the different behaviors of the preterm infant will be classified, and finally identifying a neonate with brain damage from a healthy one.

At the end, we would like to develop a second platform (based on NRP platform) that will integrate IoT connectivity technologies, the visualization functionalities of the NRP platform, and a software system to support functionalities based on the use of artificial intelligence techniques, real-time notifications, etc. with the aim to support healthcare professionals in the behavioral understanding of preterm infants and diagnosis of possible signs of neurological injuries at real time.

Stage 5: *Validation for early diagnosis*

The goal is to get a software-hardware platform with TRL5 (Level 5 of Technology Readiness Level) [27], tested in a real environment by healthcare professionals. We will integrate a system with basic functionality that can record and process the audio-video stream, and offer a possible diagnosis of neurological injury sings.

Then we are planning to use the monitoring system in the NICU of the Puerta del Mar University Hospital located in Cadiz, and later to apply a survey (based on TAM-Technology Acceptance Model) [28] to the clinicians and nurses that allows us to know their opinions about this contacless and non-invasive monitoring system, and if it has helped them when diagnosing the medical condition of a preterm infant.

4. Conclusions and Further Work

Brain injury is a frequent complication in preterms that need to be diagnosed as early as possible by medical staff in NICUs. At present, monitoring of the baby at NICUs is carried out 24/7 through empirical observation of nurses or medical staff along with the use of physiological parameters, gathered through wired-connected devices to the skin of the baby. As preterms cannot talk, clinical care include also intensive visual observation of the motor activity and aspect of the neonate with the aim to assess their health status.

There is much research through the use of audio and video data and AI-based techniques to support a non-contact monitoring of neonates with the goal of helping doctors in the early detection of brain damage, and improving the comfort of the babies in the incubators.

This paper presents an approach based on the use of cutting-edge technologies (AI, IoT, CV), the use of audio/video data sources and novel AI-based techniques to provides a support solution for medical staff to detect brain injuries in neonates at early stages of health development. The approach contains a methodology composed by five stages.

The paper describes the five stages but focus on the first stage (System deployment), presenting the first recording system prototype and NRP, a recording platform to gather information from different video and audio sources, providing also a labeling system and automatic recognition of physiological parameters.

As application example of this first, we described shortly the second stage, that is, the installation of the system in the NICU of University Hospital Puerta del Mar located in Cadiz, in order to gather the audio and video data. Likewise, we are going to keep progress on the research development in order to achieve the final goal.

Acknowledgments

Results incorporated in this publication received funding from the *EU Horizon 2020 MSCA-ITN-ETN "PremAtuRe nEwborn motor and cogNitive impairment: Early diagnosis - PARENT", GrantAgreement N° 956394.*. We would like to thank the clinicians, Isabel and Simon, of the hospitals hosting this work for their motivation, professionalism, and support to this research.

References

[1] Cabon, S., Porée, F., Simon, A., Met-Montot, B., Pladys, P., Rosec, O.,and Carrault, G. (2019). Audio-and video-based estimation of the sleep stages of newborns in Neonatal Intensive Care Unit. Biomedical Signal Processing and Control, 52, 362-370.

[2] Khanam, F. T. Z., Perera, A. G., Al-Naji, A., Gibson, K., and Chahl, J. (2021). Non-contact automatic vital signs monitoring of infants in a neonatal intensive care unit based on neural networks. Journal of Imaging, 7(8), 122.

[3] Khanam, F. T. Z., Al-Naji, A., Perera, A. G., Gibson, K., and Chahl, J. (2022). Non-contact automatic vital signs monitoring of neonates in NICU using video camera imaging. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1-8.

[4] Cabon, S., Porée, F., Simon, A., Rosec, O., Pladys, P., and Carrault, G. (2019). Video and audio processing in paediatrics: a review. Physiological measurement, 40(2), 02TR02.

[5] Villarroel, M., Guazzi, A., Jorge, J., Davis, S., Watkinson, P., Green, G. and Tarassenko, L. (2014). Continuous non-contact vital sign monitoring in neonatal intensive care unit. Healthcare technology letters, 1(3), 87-91

[6] Dargaville, P. A., Marshall, A. P., Ladlow, O. J., Bannink, C., Jayakar, R., Eastwood-Sutherland, C. and Gale, T. J. (2022). Automated control of oxygen titration in preterm infants on non-invasive respiratory support. Archives of Disease in Childhood-Fetal and Neonatal Edition, 107(1), 39-44.

[7] Nagy, Ádám, et al. "Continuous camera-based premature-infant monitoring algorithms for NICU." Applied Sciences 11.16 (2021): 7215.

[8] Bik, A., Sam, C., de Groot, E. R., Visser, S. S., Wang, X., Tataranno, M. L. and Dudink, J. (2022). A scoping review of behavioral sleep stage classification methods for preterm infants. Sleep Medicine.

[9] Sentner, T., Wang, X., de Groot, E. R., van Schaijk, L., Tataranno, M. L., Vijlbrief, D. C., and Dudink, J. (2022). The Sleep Well Baby project: an automated real-time sleep-wake state prediction algorithm in preterm infants. Sleep.

[10] Ihlen, E. A., Støen, R., Boswell, L., de Regnier, R. A., Fjørtoft, T., Gaebler-Spira, D., and Adde, L. (2019). Machine learning of infant spontaneous movements for the early prediction of cerebral palsy: A multi-site cohort study. Journal of clinical medicine, 9(1), 5.

[11] Sakkos, D., Mccay, K. D., Marcroft, C., Embleton, N. D., Chattopadhyay, S., and Ho, E. S. (2021). Identification of abnormal movements in infants: A deep neural network for body part-based prediction of cerebral palsy. IEEE Access, 9, 94281-94292.

[12] Leo, M., Bernava, G. M., Carcagnì, P., and Distante, C. (2022). Video-Based Automatic Baby Motion Analysis for Early Neurological Disorder Diagnosis: State of the Art and Future Directions. Sensors, 22(3), 866.

[13] Adde, L., Helbostad, J. L., Jensenius, A. R., Taraldsen, G., Grunewaldt, K. H., and StØen, R. (2010). Early prediction of cerebral palsy by computer-based video analysis of general movements: a feasibility study. Developmental Medicine and Child Neurology, 52(8), 773-778.

[14] Adde, L., Helbostad, J. L., Jensenius, A. R., Taraldsen, G., and Støen, R. (2009). Using computer-based video analysis in the study of fidgety movements. Early human development, 85(9), 541-547.

[15] Shin, H. I., Shin, H. I., Bang, M. S., Kim, D. K., Shin, S. H., Kim, E. K., and Lee, W. H. (2022). Deep learning-based quantitative analyses of spontaneous movements and their association with early neurological development in preterm infants. Scientific Reports, 12(1), 1-9.

[16] Liu, L., Li, Y., and Kuo, K. (2018, March). Infant cry signal detection, pattern extraction and recognition. In 2018 International Conference on Information and Computer Technologies (ICICT) (pp. 159-163). IEEE.

[17] Orlandi, S., Garcia, C. A. R., Bandini, A., Donzelli, G., and Manfredi, C. (2016). Application of pattern recognition techniques to the classification of full-term and preterm infant cry.

Journal of Voice, 30(6), 656-663.

[18] Liu, L., Li, W., Wu, X., and Zhou, B. X. (2019). Infant cry language analysis and recognition: an experimental approach. IEEE/CAA Journal of Automatica Sinica, 6(3), 778-788.

[19] Manfredi, C., Bandini, A., Melino, D., Viellevoye, R., Kalenga, M., and Orlandi, S. (2018). Automated detection and classification of basic shapes of newborn cry melody. Biomedical Signal Processing and Control, 45, 174-181

[20] Raboshchuk, G., Jančovič, P., Nadeu, C., Lilja, A. P., Köküer, M., Mahamud, B. M., and Veciana, A. R. D. (2015). Automatic detection of equipment alarms in a neonatal intensive care unit environment: A knowledge-based approach. In the Sixteenth Annual Conference of the International Speech Communication Association.

[21] Digi-NewB - GCS HUGO - CHU - monitoring system, http://www.digi-newb.eu, accessed 14 July 2022

[22] Cabon, S., Porée, F., Cuffel, G., Rosec, O., Geslin, F., Pladys, P. and Carrault, G. (2021). Voxyvi: A system for long-term audio and video acquisitions in neonatal intensive care units. Early Human Development, 153, 105303.

[23] Einspieler, C., Prechtl, H. F., Ferrari, F., Cioni, G., and Bos, A. F. (1997). The qualitative assessment of general movements in preterm, term and young infants—review of the methodology. Early human development, 50(1), 47-60.

[24] Porro, M., Fontana, C., Giannì, M. L., Pesenti, N., Boggini, T., De Carli, A.and Picciolini, O. (2020). Early detection of general movements trajectories in very low birth weight infants. Scientific Reports, 10(1), 1-7.

[25] Einspieler, C., and Prechtl, H. F. (2005). Prechtl's assessment of general movements: a diagnostic tool for the functional assessment of the young nervous system. Mental retardation and developmental disabilities research reviews, 11(1), 61-67

[26] Einspieler, C., Prechtl, H. F. R., Bos, A. F., Ferrari, F. and Cioni, G. Prechtl 's Method on the Qualitative Assessment of General Movements in Preterm, Term and Young Infants (Mac Keith Press, 2004).

[27] Martínez-Plumed, F., Gómez, E., and Hernández-Orallo, J. (2021). Futures of artificial intelligence through technology readiness levels. Telematics and Informatics, 58, 101525.

[28] Na, S., Heo, S., Han, S., Shin, Y., and Roh, Y. (2022). Acceptance Model of Artificial Intelligence (AI)-Based Technologies in Construction Firms: Applying the Technology Acceptance Model (TAM) in Combination with the Technology–Organisation–Environment (TOE) Framework. Buildings, 12(2), 90.