2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

CradlePosture: Camera-Based Approach for Estimating Neonate's Posture Based on Caregiver's Holding Behaviors

Hiyori Tsuji[1], Takumi Yamamoto[1], Maiko Kobayashi[2], Kyoshiro Sasaki[3], Noriko Aso[4], Yuta Sugiura[1]

***Abstract*— In holding behaviors, the posture of a neonate is crucial for ensuring safety. In this study, assuming a system for users to learn proper holding posture, we propose an approach for measuring the posture of a neonate during holding behaviors. Our method estimates the posture angles of a neonate based on holding behaviors captured by a camera. By using regression models to estimate a neonate's posture angles based on the posture of a person holding a neonate, estimation can be made of the posture angles of the neonate without directly attaching sensors to an actual neonate. The results show for within-participant validation, both the inclination angle– the angle at which the doll is rotated around its chest from a horizontal position towards raising its head– and the adduction angle– the angle at which the doll is rotated from a horizontal position around its body axis towards the person's body side– could be estimated with high accuracy, but for between-participant validation, only inclination angles could be estimated.**

I. INTRODUCTION

The crying of infants can negatively impact the mental health of caregivers [10], which can lead to child abuse [22]. Caregivers often engage in holding behavior to soothe crying infants. This kind of early, positive physical interaction is considered important for both caregivers and infants. For example, skin-to-skin contact between caregiver and infant immediately after childbirth is believed to have long-term effects on caregiver sensitivity and infant self-regulation [6]. Additionally, physical contact in holding behaviors also plays an essential role in the development of the child's self-regulation abilities and the formation of attachment, or bond, between caregivers and children [27].

One critical aspect of neonatal care is the posture maintained when holding neonates. Specifically, for newborns unable to support their own heads, the neck angle plays a pivotal role in facilitating unobstructed breathing, potentially preventing life-threatening situations [7]. Therefore, caregivers must master and employ master and employ holding techniques that ensure optimal positioning for newborns. However, training opportunities for acquiring such techniques are predominantly available only in professional settings, such as hospitals. This gap highlights the need for a system that can monitor and analyze the posture of newborns during holding and provide real-time feedback to caregivers. By having an accessible, user-friendly platform that offers guidance on correct holding postures, caregivers could improve their skills outside of professional environments, ensuring better neonatal care at home.

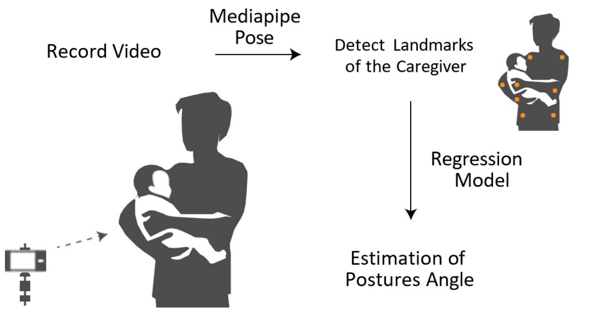

Fig. 1. Method for estimating the posture angles of a neonate based on holding behaviors captured by a camera

To develop such a system, a neonate's posture must be measured while a caregiver is holding the neonate with feedback provided for appropriate posture adjustments. For measuring posture, two approaches are considered. The first involves attaching wearable sensors directly to the neonate. Although this method simplifies measurement due to direct sensor attachment, caregiver reluctance to use wearable sensors on neonates is a significant drawback. An alternative is camera-based measurement, which eliminates the need for direct attachment and has been extensively explored in existing research. For this measurement, many frameworks are already available. These models are primarily built using data on adult posture and thus there may be limitations in the accuracy of recognition when measuring the posture of neonates, which differs significantly from adults in terms of characteristics.

In this paper, we propose a method to estimate the posture angles of a neonate based on holding behaviors captured by a camera, assuming that the future implementation of a training system uses this method when a user is holding an actual neonate (Fig. 1). We evaluated the cradle hold, which is the basic holding posture of a neonate. We focused on the posture angles that change depending on holding behaviors and are related to the safety of neonates. First, we filmed the act of holding a doll simulating a neonate

[1]Hiyori Tsuji, Takumi Yamamoto, and Yuta Sugiura are with the Faculty of Science and Engineering, Keio University, 4-1-1, Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8521, Japan `hiyori-t.01, imuka06x17, sugiura@keio.jp`

[2]Maiko Kobayashi is with the Research Institute for Science and Engineering, Waseda University, 3-4-1 Ohkubo, Shinjuku, Tokyo 169-8555, Japan `maikoyoshida417@gmail.com`

[3]Kyoshiro Sasaki is with the Faculty of Informatics, Kansai University, 2-1-1, Ryozenji-cho, Takatsuki, Osaka, 569-1095, Japan `kyoshiro0920@gmail.com`

[4]Noriko Aso is with the Faculty of Human Sciences, Kanagawa University, 3-27-1, Rokkakubashi, Kanagawa-ku, Yokohama, Kanagawa 221-8686, Japan `ft101895ru@kanagawa-u.ac.jp`

979-8-3315-3161-4/25/$31.00 ©2025 IEEE

and simultaneously measured the doll's posture angles using Xsens DOT sensor attached to the doll. MediaPipe [3], a machine learning framework developed by Google, was used to estimate the posture of a person holding a neonate, and regression models were constructed to estimate the posture angle of a neonate based on the landmark coordinates. Through user experiments, we evaluated the accuracy of the models' estimation of posture angles and found for within-participant validation both the inclination angle and adduction angle could be estimated with high accuracy, but for between-participant validation only the inclination angles could be estimated.

II. RELATED WORK

A. Childcare Support System

There has been an increase in research focusing on neonates and infants [21], [18], [24]. Among these studies, childcare support systems have been confirmed to reduce the psychological burden on caregivers [4]. For example, Estrellita is a tool designed to record and support the health status of premature children [13]. In addition, studies focusing on infant crying have proposed a model to detect infant crying in the real world [28] and a system to classify types of neonate crying and predict the reasons for crying [25].

Research has been conducted on systems to monitor the health status and manage the safety of babies using sensors. The main measurement approaches can be divided into two: attaching wearable sensors to babies and embedding sensors in the surrounding environment. First, as approaches to attach wearable sensors to infants, systems involve wearable devices like belts to obtain and monitor infants' biosignal data and provide feedback to caregivers [9], [2], [29]. Additionally, models have been proposed to detect holding behaviors from sensor values [27], and to classify children's daily activities using an accelerometer and a barometric pressure sensor attached to the waist [19].

On the other hand, research has also been conducted on measuring by attaching sensors to the environment where babies live or to the equipment used by babies [30], [12], [20], [1], [11]. Baby Lucent [12] tracks health and nutrition status using a sensor-equipped pacifier and a feeding bottle. Angelcare [1] has developed products to monitor the movement of babies with a camera and sensor pad. Doryan et al. [11] proposed a model to infer contact between parent and child from camera images.

In this study, we propose a camera-based approach to measure the posture angles of a neonate while the neonate is being carried, thereby managing the safety of the neonate. The wearable-sensor-based methods require the attachment of sensors to the child's body or devices equipped with sensors to monitor the child's condition. Considering the cost and availability, the wearable-sensor-based method might not be suitable for easy use at home, and some caregivers are reluctant to attach sensors directly to their children. In contrast, the proposed method can be applied using readily available cameras such as smartphones, and enables measurement without the need to attach a sensor. Additionally, although Doryan et al. [11] proposed a camera-based method for measuring caregiver-child interactions, they focused on the classification of contact and non-contact states. On the other hand, this study focuses specifically on holding behaviors during parent-child interactions. Thus, this research differs in its focus on estimating the posture angles of children during holding behaviors.

B. Camera-Based Posture Estimation

Studies on human posture estimation from camera images have been proposed as methods for evaluating human movements. These estimates can be used to grasp and classify postures and to suggest appropriate skill training. Wang et al.'s study [26] uses a CNN model for feature extraction and performs posture estimation based on spatial and temporal keypoint relations. Bhamidipati et al.'s study [5] uses MediaPipe for posture estimation and analyzes the posture by calculating angles between landmarks with OpenCV. Furthermore, AI Golf [17] detects minor differences between professional and user movements, generates intermediate steps, and helps users learn the ideal form gradually. However, these posture estimation models are constructed using data from adult postures, and therefore have limitations when it comes to applying them to neonates, who have different appearances and postures.

Some research has been conducted to estimate infant postures from videos. Stahl et al. extracted motion information from optical flow [23]. When using optical flow, high frame rates are required to capture frame-by-frame differences, and careful parameter selection is necessary for accurate tracking. As a method that does not use optical flow, Hesse et al. proposed a model that captures infant motion and shape from RGB-D data [14]. However, the need for an RGB-D camera makes it difficult for each home to obtain one. Chambers et al. extended pose-estimation algorithm, OpenPose, to handle infant postures [8]. Nevertheless, this model estimates the posture of an infant lying on its back, and requires camera placement that captures the entire body within the frame. Therefore, it is challenging to estimate the posture of an infant being held.

Moreover, research has been conducted to estimate the values of wearable sensors from camera video. In studies of estimation with MediaPipe, Leddy et al. [16] assured the concurrent validity of MediaPipe posture estimation and IMU system, and reported a strong correlation. IMUTube [15] is a pipeline that estimates IMU sensor values based on video for the purpose of Human Activity Recognition.

Based on these studies, our research conducts posture estimation of a person performing holding behaviors with MediaPipe. At the same time, by measuring the posture angles of the target with Xsens Dot sensor, regression models are created to estimate the posture angles of the target to be soothed from the posture of the person.

III. MATERIAL AND METHOD

In this paper, we propose a method using a camera to estimate the posture angles of a soothed neonate based on

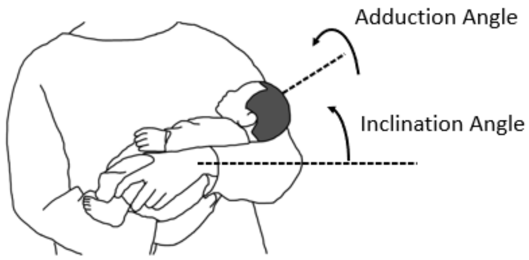

Fig. 2. Definition of inclination angle and adduction angle

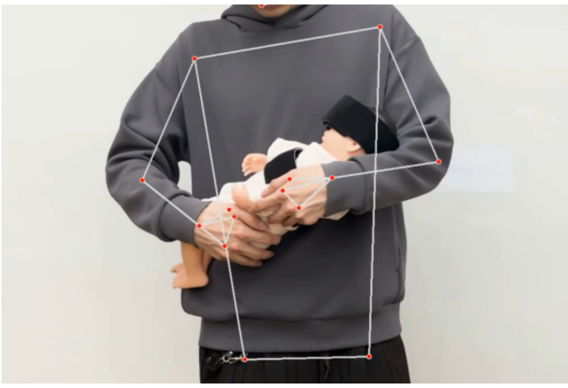

Fig. 3. Posture estimation with MediaPipePose

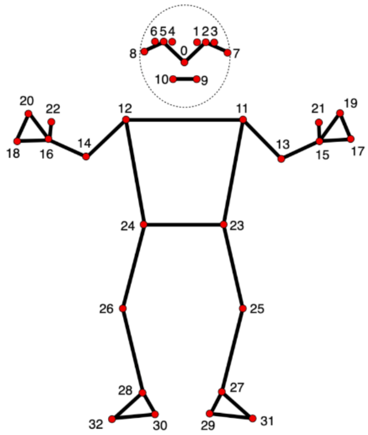

Fig. 4. Landmark location and number

the posture of a caregiver holding the neonate. The following sections describe (1) dataset acquisition and preprocessing, and (2) feature extraction and regression models construction based on the dataset.

A. Dataset Acquisition and Preprocessing

To estimate the posture angles of a neonate during the holding behaviors from camera images, a dataset is needed that corresponds to images of the holding postures and their corresponding true values of the posture angles. A user experiment was conducted to acquire the dataset. In this paper, we focused on the inclination angle and adduction angles as posture angles. As shown in Fig. 2, each angle was defined as follows: the inclination angle was the angle at which the doll was rotated around its chest from a horizontal position towards raising its head, and the adduction angle was the angle at which the doll was rotated from a horizontal position around its body axis towards the person's body side. This experiment was approved for ethics application at the author's home institution.

The participants included three males and two females (mean age = 22.6 years; SD = 0.5). Using a smartphone camera, we captured videos from the front of the participants holding the doll, ensuring sufficient distance to capture each participant from head to waist. Simultaneously, we measured the doll's posture angles using Xsens Dot sensor attached to the doll's chest. The reason for capturing the upper body is that we believe the movements of the caregiver's arms, shoulders, and waist significantly influence the posture angles of the neonate. Participants were instructed to perform soothing motions twice with a cradle hold (with the doll's head on the left), each session lasting approximately 30 seconds. They were asked to gradually increase or decrease the inclination angle during the first session and the adduction angle during the second session, with changes made every 10 seconds. However, no instructions were given regarding the adduction angle during the first session and the inclination angle during the second session. Videos were recorded at 60 fps using the standard camera application on a smartphone, and the Xsens Dot sensor measured at 60 Hz. These data were synchronized manually. We constructed separate regression models for the inclination angle and the adduction angle estimations. The inclination angle data measured by the Xsens Dot sensor in the first session was used as the objective variable for the model that estimates the inclination angle, and the adduction angle data measured in the second session was used as the objective variable for the model that estimates the adduction angle.

The distribution of the number of data frames used for training was aligned for each participant due to variations in the distribution of truth values. Based on these data, 140 frames for the inclination angle and adduction angle were randomly selected in the angle ranges of every 10 degrees within the angle range of 0 to 40 degrees.

B. Feature Extraction and Regression Models Construction

In this study, regression models were constructed to estimate the posture angles of a neonate based on the videos of neonate holding behaviors. First, for feature extraction from the images, we extracted the postures of the participants performing the holding behaviors, inferred as shown in Fig. 3. For human posture estimation from the video, we used MediaPipePose [3], a machine-learning framework developed by Google. By inputting the video into MediaPipePose, the positional coordinates of landmarks on the entire human body can be detected. After extracting the 25 absolute landmark coordinates of the upper body, the coordinates were converted to relative coordinates with the origin at the midpoint of the participant's shoulder coordinates to

TABLE I

WITHIN-PARTICIPANT VALIDATION RESULTS FOR INCLINATION AND ADDUCTION ANGLE ESTIMATION MODELS

		P1	P2	P3	P4	P5	Mean
Inclination angle estimation model	R^2	0.957	0.969	0.969	0.944	0.965	0.961±0.009
	MAE	0.97	0.76	0.71	0.96	0.79	0.839 ±0.106
Adduction angle estimation model	R^2	0.688	0.973	0.926	0.930	0.972	0.898±0.107
	MAE	2.58	0.72	1.04	1.22	0.683	1.250 ±0.696

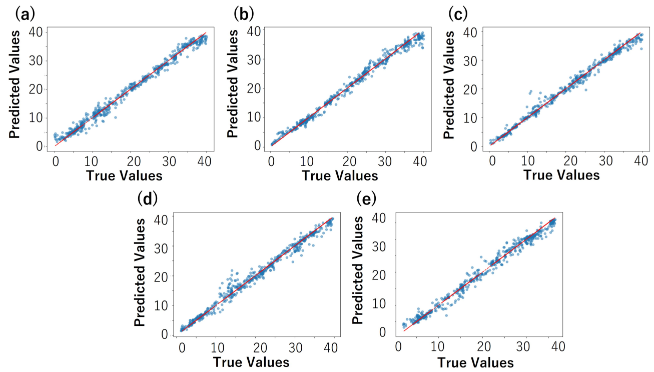

Fig. 5. Correspondence between the true values and the predicted values in the inclination angle estimation model for (a) P1, (b) P2, (c) P3, (d) P4, and (e) P5

account for changes in coordinate values caused by their movements and camera shake (Fig. 4 Landmark 0-24). We used the relative coordinates as the explanatory variables and constructed the regression models to estimate the doll's posture angles, the inclination angle and the adduction angle. We employed a random forest regressor to construct the models. The random forest model was implemented using Python version 3.11.3 on a Windows 11 operating system. The model was trained with the following hyperparameters: the number of trees (n_estimators) = 64, the number of features considered for splitting at each node (max_features) = one-third of the total number of features, and the random seed (random_state) = 0. All other hyperparameters were set to their default values of sklearn.

C. Evaluation Method

To examine the influence of the individual differences in the models, we conducted both within-participant and between-participant validations. For the within-participant validation, we developed models for each of the five participants and performed 5-fold cross-validation. For the between-person validation, we conducted cross-validation where one person's data was set aside as the test dataset while the data from the remaining four participants were used for training. The evaluation indicators used were Mean Absolute Error (MAE) and coefficient of determination(R^2).

IV. RESULT AND ANALYSIS

A. Result

FIve-fold cross-validation was conducted to evaluate the regression models for each participant, and Table I shows the results for R^2, MAE and their respective standard deviations.

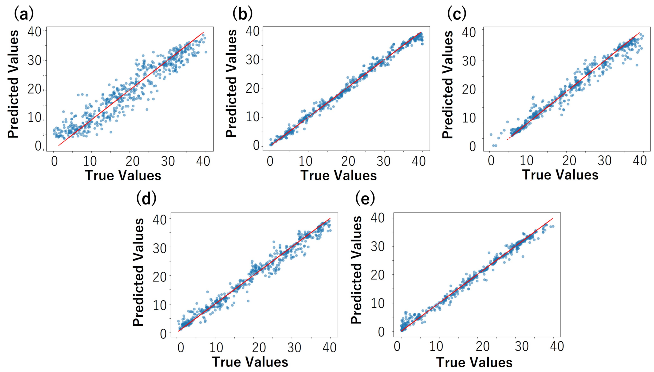

Fig. 6. Correspondence between the true values and the predicted values in the adduction angle estimation model for (a) P1, (b) P2, (c) P3, (d) P4, and (e) P5

Additionally, Fig. 5 and Fig. 6 show the correspondence between the true values and the predicted values for each model. However, considering that the coefficient of determination increased with more explanatory variables in a multiple regression analysis, we used an adjusted coefficient of determination.

For the inclination angle estimation models, the mean coefficients of determination and MAE were 0.961 ± 0.009 and 0.839 ± 0.106. In contrast, for the adduction angle estimation models, these values were 0.898 ± 0.107 and 1.250 ± 0.696, respectively. The coefficients of determination were high for all models except for the adduction angle estimation model of participant P1. The MAE was within 1 degree for the inclination angle estimation model and within 3 degrees for the adduction angle estimation model, both of which represented minimal errors in capturing human movements. The standard deviations of both R^2 and MAE were smaller for the inclination angle estimation model.

Table II and Fig. 7 show the results of R^2, MAE, and their respective standard deviations for each model in the between-participant validation, as well as the correspondence between the true values and predicted values.

For the inclination angle estimation model and the adduction angle estimation model, the coefficients of determination values were 0.625 ± 0.003 and -0.206 ± 0.513, respectively. This indicates that while the inclination angle estimation model worked, the adduction angle estimation model did not. Compared with the results from within-participant validation, the MAE for both models were also larger. Additionally, the standard deviation of each indicator was smaller for the inclination angle estimation model, similar to the within-participant validation.

B. Analysis

The within-participant validation demonstrates that the estimation models achieved high accuracy, with coefficients of determination of 0.9 or higher for all participants' data, except for the P1 adduction angle estimation model. The mean and standard deviation of the evaluation indicators for each participant's estimation model show that the inclina-

TABLE II
EVALUATION FOR EACH MODEL

	Inclination angle	Adduction angle
R^2	0.625±0.003	-0.206±0.513
MAE	5.12±0.29	8.69±2.96

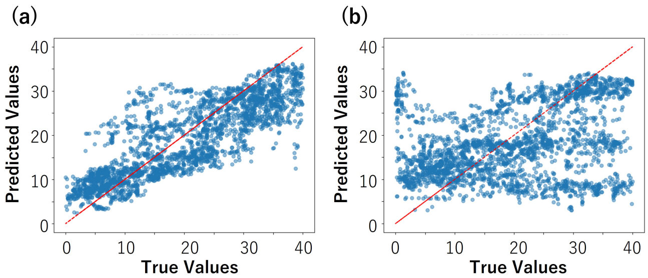

Fig. 7. Correspondence between true values and predicted values in the estimation model for (a) inclination angle and (b) abduction angle

tion angle estimation was more accurate and exhibited less variation in accuracy. This can be attributed to the ease of capturing changes in the inclination angle from the videos. As the inclination angle changes, the height of the left elbow also changes significantly. Since the videos were captured from the front of the participant, these changes could be clearly observed in the video. In contrast, changes in the adduction angle represent depth information and are difficult to capture with a camera from the front. We believe that this caused the low estimation accuracy of the adduction angle. However, since the participant's own data for both training and testing were used to construct the models, we were able to capture individual patterns of motion for each participant. Thus, the results for both the inclination angle and the adduction angle were highly accurate.

On the other hand, for the between-participant validation, we observed that while the inclination angle could be estimated, the adduction angle could not. The significant decrease in accuracy for adduction angle estimation can be attributed to the challenge of capturing depth information from the camera angle. We also consider that the estimation models were not universally applicable to all participants due to variations in how each participant performed the motion.

V. DISCUSSION, LIMITATIONS AND FUTURE WORK

A. Experimental Design

To construct the regression models, we extracted 140 frames at intervals of 10 degrees within the range of 0 to 40 degrees for training and testing data. This was done to ensure the data were evenly distributed among all participants. However, we believe a wider range of angles is necessary for assessing whether the posture is appropriate. Additionally, even within each interval of 10 degrees, there may be biases in the angles, which could affect the accuracy. Therefore, the movement instructions must be clarified regarding angles during data collection and to ensure that participants' movements are as consistent as possible.

In addition, in the experiment conducted to obtain sensor data, the holding method was specified as a cradle hold. However, there are various ways to hold babies, including carrying them under the arms and upright, etc. In future studies, we plan to collect sensor data from different holding positions.

For the between-participant validation, the accuracy was low, particularly in the estimation of the adduction angle, which was difficult to capture with a front-facing camera. This suggests that information from the front alone was insufficient. Changes in the adduction angle are associated with the bending of the arms towards the body, indicating the need for additional information from side-view cameras. Using multiple cameras is also expected to help mitigate the occlusion problem, where objects are obstructed by other objects. While increasing the number of cameras may improve accuracy, the increase in the number of explanatory variables could lead to over-fitting, heavy computational costs, and a more complicated interpretation of the models. Therefore, careful consideration should be given to the appropriate number of cameras and camera angles.

B. For Practical Application

In this experiment, recruiting actual neonates, was difficult so we used a doll instead to simulate a neonate. For this reason, we did not validate whether the method is applicable when a user holds an actual neonate. There is a possibility of significant errors in posture estimation of the person who is holding a neonate due to the large number of movements of the neonate. In the future, we will collect sensor data when holding an actual neonate.

In this paper, we propose a method for evaluating holding behaviors and assume the creation of a system that enables users who have just started child-rearing to train their holding behaviors at home using this method. However, achieving highly accurate estimation with the current regression models requires users to collect their own data first, which is time-consuming. Thus, estimation accuracy for between-participant validation must be enhanced and made applicable to untrained users, thus reducing the user burden. In the future, we will improve the machine learning model to improve accuracy.

Regarding feedback methods in the system, real-time feedback on the screen and through audio based on the user's movements can be considered, as in the case of Pababy [18] and Tsuji et al. 's research [24]. Evaluating the user and suggesting improvements using images and videos would also be beneficial, as in the studies AI Coach [26] and AI Golf [17]. As future work, we aim to develop an application utilizing this system and investigate its effectiveness.

VI. CONCLUSION

This paper proposes a method for measuring the body posture of a neonate based on a user's holding behaviors. The method estimates the posture angles of a neonate from the holding behaviors captured by a camera. Based on this method, we are going to create a system that provides feedback to ensure that a neonate being held conforms to appropriate body postures, which enables caregivers to train their holding postures at home. Future plans include improving the accuracy of the estimation for between-participant validation.

ACKNOWLEDGMENT

Part of this work was supported by JSPS KAKENHI (Grant Number JP23H01046).

REFERENCES

[1] Angelcare – angelcarebaby. https://angelcarebaby.com/. (Accessed on 03/22/2024).

[2] Home — temptraq. https://temptraq.healthcare/. (Accessed on 03/22/2024).

[3] Pose landmark detection guide — mediapipe — google for developers. https://developers.google.com/mediapipe/solutions/vision/pose_landmarker. (Accessed on 01/16/2024).

[4] Lesley Barclay, Louise Everitt, Frances Rogan, Virginia Schmied, and Aileen Wyllie. Becoming a mother—an analysis of women's experience of early motherhood. *Journal of advanced nursing*, 25(4):719–728, 1997.

[5] Venkata Sai P Bhamidipati, Ishi Saxena, D. Saisanthiya, and Mervin Retnadhas. Robust intelligent posture estimation for an ai gym trainer using mediapipe and opencv. In *2023 International Conference on Networking and Communications (ICNWC)*, pages 1–7, 2023.

[6] Ksenia Bystrova, Valentina Ivanova, Maigun Edhborg, Ann-Sofi Matthiesen, Anna-Berit Ransjö-Arvidson, Rifkat Mukhamedrakhimov, Kerstin Uvnäs-Moberg, and Ann-Marie Widström. Early contact versus separation: effects on mother–infant interaction one year later. *Birth*, 36(2):97–109, 2009.

[7] Mary E. Case. Abusive head injuries in infants and young children. *Legal Medicine*, 9(2):83–87, 2007. Special section: Forensic Pathology: New Concepts and Technologies.

[8] Claire Chambers, Nidhi Seethapathi, Rachit Saluja, Helen Loeb, Samuel R. Pierce, Daniel K. Bogen, Laura Prosser, Michelle J. Johnson, and Konrad P. Kording. Computer vision to automatically assess infant neuromotor risk. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(11):2431–2442, 2020.

[9] Wei Chen, Sietse Dols, Sidarto Bambang Oetomo, and Loe Feijs. Monitoring body temperature of newborn infants at neonatal intensive care units using wearable sensors. In *Proceedings of the Fifth International Conference on Body Area Networks*, BodyNets '10, page 188–194, New York, NY, USA, 2010. Association for Computing Machinery.

[10] Kaya de Barbaro, Megan Micheletti, Xuewen Yao, Priyanka Khante, Mckensey Johnson, and Sherryl Goodman. Infant crying predicts real-time fluctuations in maternal mental health in ecologically valid home settings. *Developmental Psychology*, 2023.

[11] Metehan Doyran, Ronald Poppe, and Albert Ali Salah. Embracing contact: Detecting parent-infant interactions. In *Proceedings of the 25th International Conference on Multimodal Interaction*, ICMI '23, page 198–206, New York, NY, USA, 2023. Association for Computing Machinery.

[12] Kevin Gaunt, Júlia Nacsa, and Marcel Penz. Baby lucent: pitfalls of applying quantified self to baby products. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, page 263–268, New York, NY, USA, 2014. Association for Computing Machinery.

[13] Gillian R. Hayes, Karen G. Cheng, Sen H. Hirano, Karen P. Tang, Marni S. Nagel, and Dianne E. Baker. Estrellita: A mobile capture and access tool for the support of preterm infants and their caregivers. *ACM Trans. Comput.-Hum. Interact.*, 21(3), jun 2014.

[14] Nikolas Hesse, Sergi Pujades, Javier Romero, Michael J. Black, Christoph Bodensteiner, Michael Arens, Ulrich G. Hofmann, Uta Tacke, Mijna Hadders-Algra, Raphael Weinberger, Wolfgang Müller-Felber, and A. Sebastian Schroeder. Learning an infant body model from rgb-d data for accurate full body motion analysis. In Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 792–800, Cham, 2018. Springer International Publishing.

[15] Hyeokhyen Kwon, Catherine Tong, Harish Haresamudram, Yan Gao, Gregory D. Abowd, Nicholas D. Lane, and Thomas Plötz. Imutube: Automatic extraction of virtual on-body accelerometry from video for human activity recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(3), sep 2020.

[16] Chloe Leddy, Richard Bolger, Paul J Byrne, Sharon Kinsella, and Lilibeth Zambrano. Concurrent validity of the human pose estimation model "mediapipe pose" and the xsens inertial measuring system for knee flexion and extension analysis during hurling sport motion. In *2023 IEEE International Workshop on Sport, Technology and Research (STAR)*, pages 49–52, 2023.

[17] Chen-Chieh Liao, Dong-Hyun Hwang, and Hideki Koike. Ai golf: Golf swing analysis tool for self-training. *IEEE Access*, 10:106286–106295, 2022.

[18] Jianhui Liu, Yijun Zhao, Jiadi Wang, Chenyi Xie, Fangyuan Cheng, Jintao Nie, Ge Yan, Cheng Yao, Leijing Zhou, Preben Hansen, Fangtian Ying, and Guanyun Wang. Pababy: An interactive system for first-time parents to learn neonatal nursing. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA '22, New York, NY, USA, 2022. Association for Computing Machinery.

[19] Yunyoung Nam and Jung Wook Park. Child activity recognition based on cooperative fusion model of a triaxial accelerometer and a barometric pressure sensor. *IEEE Journal of Biomedical and Health Informatics*, 17(2):420–426, 2013.

[20] Geeyoung Noh, Jooyoung Shin, Jundong Cho, and Andrea Bianchi. Guidelines for a baby's nutrition monitoring system. In *Proceedings of HCI Korea*, HCIK '15, page 92–97, Seoul, KOR, 2014. Hanbit Media, Inc.

[21] Alena Petrasova, Gabriela Czanner, Jassim Happa, Silvester Czanner, Dieter Wolke, and Alan Chalmers. Assessing a virtual baby feeding training system. In *Proceedings of the 7th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, AFRIGRAPH '10, page 37–44, New York, NY, USA, 2010. Association for Computing Machinery.

[22] Sijmen A Reijneveld, Marcel F van der Wal, Emily Brugman, Remy A Hira Sing, and S Pauline Verloove-Vanhorick. Infant crying and abuse. *The Lancet*, 364(9442):1340–1342, 2004.

[23] Annette Stahl, Christian Schellewald, Øyvind Stavdahl, Ole Morten Aamo, Lars Adde, and Harald Kirkerod. An optical flow-based method to predict infantile cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(4):605–614, 2012.

[24] Hiyori Tsuji, Takumi Yamamoto, Sora Yamaji, Maiko Kobayashi, Kyoshiro Sasaki, Noriko Aso, and Yuta Sugiura. Smartphone-based teaching system for neonate soothing motions. In *2024 IEEE/SICE International Symposium on System Integration (SII)*, pages 178–183, 2024.

[25] V. Vaishnavi and P. Suveetha Dhanaselvam. Neonatal cry signal prediction and classification via dense convolution neural network. *J. Intell. Fuzzy Syst.*, 42(6):6103–6116, jan 2022.

[26] Jianbo Wang, Kai Qiu, Houwen Peng, Jianlong Fu, and Jianke Zhu. Ai coach: Deep human pose estimation and analysis for personalized athletic training assistance. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM '19, page 374–382, New York, NY, USA, 2019. Association for Computing Machinery.

[27] X. Yao, T. Plötz, M. Johnson, and K. de Barbaro. Automated detection of infant holding using wearable sensing: Implications for developmental science and intervention. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(2), jun 2019.

[28] Xuewen Yao, Megan Micheletti, Mckensey Johnson, Edison Thomaz, and Kaya de Barbaro. Infant crying detection in real-world environments. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 131–135, 2022.

[29] Lingyan Zhang, Lijuan Liu, Weiqiang Ying, Muling Huang, Cheng Yao, and Fangtian Ying. Modhera: A modular kit for parents to take care babies. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference*, IDC '21, page 547–551, New York, NY, USA, 2021. Association for Computing Machinery.

[30] Yijun Zhao, Yong Shen, Xiaoqing Wang, Jiacheng Cao, Shang Xia, Fangtian Ying, and Guanyun Wang. Pneumat: Pneumatic interaction system for infant sleep safety using shape-changing interfaces. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21, New York, NY, USA, 2021. Association for Computing Machinery.