A Study of At-term and Preterm Infants' Motion Based on Markerless Video Analysis

Luca Garello*
Italian Institute of Technology (IIT)
MaLGa Center - DIBRIS
University of Genova
Genova, Italy
luca.garello@iit.it

Matteo Moro*
MaLGa Center - DIBRIS
University of Genova
Genova, Italy
matteo.moro@edu.unige.it

Chiara Tacchino
Istituto Giannina Gaslini
Genova, Italy
chiaratacchino@gaslini.org

Francesca Campone
Istituto Giannina Gaslini
Genova, Italy
francescacampone@gaslini.org

Paola Durand
Istituto Giannina Gaslini
Genova, Italy
paoladurand@gaslini.org

Isabella Blanchi
Istituto Giannina Gaslini
Genova, Italy
isablanchi@gaslini.org

Paolo Moretti
Istituto Giannina Gaslini
Genova, Italy
paolomoretti@gaslini.org

Maura Casadio§
DIBRIS
University of Genova
Genova, Italy
maura.casadio@unige.it

Francesca Odone§
MaLGa Center - DIBRIS
University of Genova
Genova, Italy
francesca.odone@unige.it

***Abstract*—Preterm birth is sometimes associated with neurological disorders caused by lesions of the developing brain. A diagnosis in the first weeks of child's life is important to plan timely and appropriate rehabilitative interventions for infants at risk of neuro-motor disabilities. A largely adopted method for the early diagnosis of neuro-motor disorders is the General Movements assessment, based on the evaluation of infants' spontaneous motor patterns. However, an accurate clinical assessment of infant motion requires highly specialized personnel, not always available at all sites. To insure an objective motion analysis, several studies proposed the use of marker-based techniques. Unfortunately, markers are uncomfortable and can affect the naturalness of the motion. Therefore, much effort has been dedicated in developing marker-less techniques targeting unobtrusive and reliable motion analysis. In this work we propose a marker-less video-based methodology to analyze infants' spontaneous movements in RGB videos. First, we detect relevant landmarks on the infants' body. Then, we compute kinematic parameters that describe infants' motion patterns. We validate the effectiveness of the computed parameters on a dataset of 68 infants, 27 of which with a clinically-assessed evidence of neuro-motor disorders: our method successfully discriminates infants with and without motor disorders with an accuracy of 78.2%.**

***Index Terms*—Human Motion Analysis, Video Analysis, Markerless, Semantic Features**

I. Introduction

According to the World Health Organization (WHO), preterm birth is the leading cause of death in children younger than 5 years worldwide. Preterm survival rates have increased in high-income countries thanks to the development of intensive care techniques. However, with the increase of preterm survival rate, also the incidence of neurological diseases grows and many of the surviving infants face a life of disability, including learning, visual and motor difficulties [2]. Common neurological disorders that could occur in this early stage of life can be grouped under the 'umbrella' term of Cerebral Palsy (CP): they include permanent but not progressive lesions of the developing brain [11].

An early diagnosis of pathological cases would allow the start of early rehabilitation treatment that could significantly increase the chances of recovery. In this scenario, it is crucial to find reliable and objective techniques which may support physicians in the study of children's neurological status. With the use of magnetic resonance imaging (MRI) it is possible to obtain a thorough investigation of any lesions in specific areas of the brain. However, MRI is not always available due to the high costs and to the organizational difficulties [7]. Other approaches to neurological evaluation include: (1) traditional neurological examination [15] and (2) neurological examination based on the observation of spontaneous motor behavior such as the General Movements (GM) theory [16]. GMs are spontaneous movements of variable amplitude and speed involving all parts of the body and they accurately reflect the state of neuro-motor development [16]. This type of study involves highly specialized personnel for a long amount of time, and is often operator dependent [1]. For these reasons, there is a need for objective methodologies able to extract quantitative parameters that represent infants' motion patterns.

In the field of human motion understanding, an accurate quantitative analysis of the movement can be achieved thanks to marker-based motion capture systems [13]. However, markers and sensors placed on the body are cumbersome and they could prevent the naturalness of the motion [6], especially in infants. For these reasons, recently, marker-less techniques for human motion analysis based on computer vision have been studied [5], [8]: they have the potential to solve or reduce some of the issues of marker-based approaches, as they allow for a natural person-friendly interaction, they are non invasive

* These authors contribute equally to the work
§ These authors contribute equally to the work

The authors thank Pietro Morasso, Psiche Giannoni and Marco Massimo Fato for their support to the study.

and inexpensive.

The first approaches to video-based infants' motion analysis [1], [20] are based on change detection. In both cases the parameters extracted were adopted to perform a classification task between infants with and without neuro-motors disorders. Baccinelli et al. [3] provide a semi-automatic software package with a graphic user interface in order to track the movement of hands and feet. Kawashima et al. [10] propose a method to study infants' crawling starting from videos acquired from two different view-points and extracting infant's shape-based parameters starting from a background subtraction. To the best of our knowledge there is a limited number of studies in the literature focused on markerless joint detection techniques to study infants' motion. Our work tries to address this problem. The main software requirements of the application field are a good accuracy and some level of interpretability of the results. The latter drives us towards a two-levels architecture (representation + classification) instead of an end-to-end one. We address accuracy by relying on few stable semantic features.

We propose a novel 2D marker-less motion analysis system based on a single RGB camera (see Figure 1) to evaluate infants' motion during short acquisition sessions taking place in their first weeks of life. Our goal is to identify early signs of neurological disorders. We focus on a 2D pipeline because our long term goal is to provide an easy to use methodology which could also be applied at home.

The pipeline we propose includes (1) *video representation:* we detect the (x, y) positions of relevant landmark points (nose, hands and feet) in the image plane; we filter them in order to add a spatio-temporal consistency, and then we compute quantitative parameters inspired by the neuro-motor literature to describe infant's motion [13]; (2) *classification:* the extracted parameters form a feature vector we use as input of a binary classifier for the discrimination of infants *with and without neuro-motor disorders*. This stage also incorporate a greedy feature selection procedure, to select the most meaningful parameters in the feature vector. Our experimental assessment is based on a clinical dataset that includes videos of 68 infants, with 55 born preterm, of which 27 present neuro-motor disorders. The main result of our analysis is a binary classifier able to discriminate between infants with and without neuromotor disorders.

We obtain very promising results (78.2% overall accuracy) by analysing videos acquired at the 40th gestational week; notice the clinical assessment we use as a ground truth has been reached two years later in the infants' life . This observation speaks in favour of the *predictive potential of our approach and its applicability as an early diagnosis tool*.

II. Dataset and Proposed Pipeline

A. Dataset

The dataset adopted for the study includes 68 infants. The study and the consent form signed by parents were approved by the Giannina Gaslini Hospital Institutional review board on 20/06/2013 (protocol number: IGGPM01). For each infant, one video was acquired at 40 weeks of gestational age (40 weeks after the conception) with a RGB camera (Canon Legria HF R37, acquiring at 25 frames per second (fps) with a resolution of 1080x1920 pixels). The camera was installed on a support above the child at a distance so that all the movements are always within the field of the camera itself (Figure 1). The videos present different backgrounds (and cots with different colors). Each video lasts about 5 minutes during which the infants are moving freely, supine, and they are not crying: 13 of the infants acquired are born at-term and the other 55 are born preterm. According to the medical experts' evaluations, among the preterm infants, 28 present normal motion patterns and the other 27 present neuro-motor disorders. The clinical evaluation is based on the Bayley test [4], performed by expert physicians involved in the study two years after the video recording.

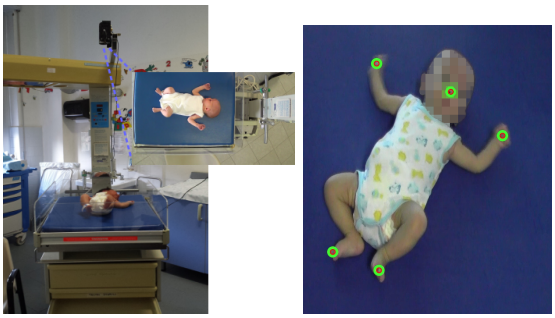

Fig. 1. Left: data acquisition setup with the camera view-point. Right: example of detected points (nose, hands, feet) - image cropped to improve visibility, face anonymised for privacy.

B. Anatomical points detection and filtering

In order to perform our analysis, we detect the position in the image plane of some anatomical points: for this study we consider few but highly distinctive semantic features (nose, hands and feet), which are stable across different subjects and acquisition conditions. For this purpose, we train a semantic feature detection architecture, DeepLabCut (DLC) [12], with examples from our dataset on the chosen anatomical points. DeepLabCut is composed by a variant of Residual Deep Network (ResNet-50) with a final deconvolutional layer to extract spatial density probability maps associated with each point. The model is pretrained on ImageNet. To train the network, we randomly select 10 frames in 55 videos (for a total of 550 frames) and we manually label the points of interest. The parameters used to train the network are the one suggested in [12] and in other applications like [14].

For each frame of each video, DeepLabCut returns a set of points $\{(x_i^t, y_i^t, \ell_i^t)\}_{t=0}^T$ with $i = \{$nose, left hand, right hand, left foot, right foot$\}$, (x_i^t, y_i^t) is the position of the $i-$th point in the $t-$th frame and ℓ_i^t —a number in the interval $[0, 1]$— is its corresponding likelihood; the latter allow us to quantify the uncertainty behind the detection of each point in each frame (see an example in Fig. 1 – green circles are the detection, red dots the ground truth). The detected features are quite stable across training and test; Table I reports the mean error (euclidean

distance) and the correspondent standard deviation (SD) in pixels for each detected point.

The choice of DeepLabCut has been mainly guided by three reasons: (1) it requires a very limited number of annotated examples for the training phase - in the order of few hundreds [12]; (2) it allows to focus only on the anatomical points we are interested in, guaranteeing a higher per-point accuracy; (3) as mentioned in [9], classical full pose estimation methods (e.g. [5]) are not always appropriate for infants' motion analysis and would require a significant amount of fine tuning and the need of a large dataset we do not possess. Moreover, DeepLabCut has the appropriate robustness-complexity trade-off for our specific study-case, where the subject movement is unconstrained, but limited by the specific development stage, the background is uniform, and only one subject is present in the scene.

To confirm this, in Table I we report the mean error in pixels for the same landmark points obtained with DeepLabCut and the full pose estimator Openpose [5].

TABLE I
MEAN ERROR ± STANDARD DEVIATION (SD) FOR EACH POINT IN PIXELS COMPUTED CONSIDERING A MANUALLY LABELED GROUND TRUTH IN 680 IMAGES.

Point	***DeepLabCut***	***Openpose***
Nose	3.73 ± 2.43	7.68 ± 3.14
Right Hand	5.12 ± 3.35	8.73 ± 3.97
Left Hand	5.34 ± 3.58	8.82 ± 4.73
Right Foot	6.57 ± 4.13	9.74 ± 5.04
Left Foot	6.18 ± 4.28	9.98 ± 5.01

To improve the stability across time of the estimated points and reduce localization errors, we proceed with a temporal processing. In particular, we need to: *(i)* correct the **mispredictions** of DeepLabCut, which occasionally detects points in a wrong position; *(ii)* manage **occlusions**, for instance if a hand is hidden by the head.

Errors due to mispredictions are easily recognizable because they involve a characteristic spike in the sequence of the point's coordinates. It is possible to overcome this problem with a median filter applied to the time sequences of the individual positions, $\{x_i\}_{i=0}^{T}$ and $\{y_i\}_{i=0}^{T}$ respectively. In the case of occlusions, the neural network will find it hard to identify the position of the occluded point as it is hidden; this situation is easy to identify as the detection likelihood ℓ_i of the occluded points drops to values close to zero. To overcome this problem and control information loss, we drop the points with a small likelihood and then, if the information loss lasts less than 2 seconds, we interpolate the trajectories in order to reconstruct the movement of each point in the temporal interval between their occlusion and their reappearance.

C. Motion descriptors

The qualitative factors that are considered by physicians during a visual motor evaluation can be summarized in (1) variability, (2) smoothness and (3) complexity of the motion [16], [17]. This knowledge is considered as a starting point to select the correspondent quantitative parameters that are computed in order to perform a computer-based evaluation of the motion. Following the considerations done in [13], we identify a correspondence between qualitative and quantitative parameters: (1) variability - cross-correlation, (2) smoothness - skewness and area out of standard deviation, (3) complexity - periodicity. In marker-based approaches, these parameters are computed on the (X, Y, Z) coordinates of markers. In our case we perform a 2D analysis of points extracted from the videos.

*1) **Cross-correlation**:* According to GM theory, it is important to determine whether the movements of upper and lower limbs are correlated [16]. To this purpose it is common to focus on the correlation between limbs' motion. We compute the correlation between limbs' speed, between their acceleration and their jerk. The same parameters are calculated for upper and lower limbs [13]. For instance, the *cross correlation* (X_{corr}) between the speed (v) of left (L) and right (R) upper/lower limb on a window of N frames (t is the index for each frame) can be written as:

$$Xcorr_v = \frac{\sigma^2_{v,L-R}}{\sigma_{v,L} * \sigma_{v,R}} \tag{1}$$

with $\sigma_{v,L}$ and $\sigma_{v,R}$ the standard deviation of the right and left speed profile and with

$$\sigma^2_{v,L-R} = \frac{1}{N-1} * \sum_{t=1}^{N-1} (v_{L,t} - \overline{v}_L) * (v_{R,t} - \overline{v}_R) \tag{2}$$

$\overline{v}$ the mean speed across all the frames.

*2) **Skewness**:* The skewness is a statistical parameter that allows us to study the distribution of speed (v) of the upper and lower limbs [13], [19]. This parameter is useful because, on average, pathological cases have peaks of higher speed. For a time window of N frames (with t as index for each frame), it is defined as:

$$Skewness(v) = \frac{\frac{1}{N-1} * \sum_{t=1}^{N-1} (v_t - \overline{v})^3}{\sigma_v^3} \tag{3}$$

with $\overline{v}$ the mean speed computed across all the frames and σ_v the correspondent standard deviation.

*3) **Area out of standard deviation of moving average and area differing from moving average**:* An important aspect evaluated by clinicians is movements smoothness. We consider the area differing from moving average as a measure that detects the divergence between the real trajectory of a point belonging to a limb and the trajectory's moving average of the same point.

If we consider the x coordinate of a certain point, the moving average $\tilde{x}$ is the mean of the values of the x coordinate in each t-th time instant for a total number of k frames is $\tilde{x} = \frac{1}{k} \sum_{t=j-\frac{k-1}{2}}^{j+\frac{k-1}{2}} x_t$ where $j = \frac{k+1}{2}, \ldots, N - \frac{k-1}{2}$ represents the index that allows the selection of k consecutive frames over their total number N (in our case k = 99, that is equivalent to about 4 seconds). Then the differences between the moving average $\tilde{x}$ and the real trajectory x are summed for all the k frames in the time window:

$$A_x = \sum_{j=\frac{k+1}{2}}^{N-\frac{k-1}{2}} |x_j - \tilde{x}_j| \quad (4)$$

Normalizing the differences between moving average and trajectory on measurement length it is possible to obtain $A_{norm,x} = \frac{A_x}{N-k}$. Finally the calculated areas of each spatial axes (x and y) and for both left (L) and right (R) limb are merged $A = \sum_s A_{norm,s,L} + \sum_s A_{norm,s,R}$, with $s = [x, y]$. In the same way, taking in account $\tilde{x}$ and its standard deviation it is possible to compute the area out of standard deviation of moving average.

*4) **Periodicity**:* This parameter is important to have an idea of the degree of complexity of the motion. To compute the periodicity of the trajectory of a certain landmark point, the number of intersections between the trajectory and its mean is determined in a fixed window of frames for each coordinate (x, y). Then, the temporal distance d_i between two intersections is computed and expressed in number of frames. The mean μ_d and standard deviation σ_d of the distances d_i are calculated, and the periodicity P (e.g. for the x coordinate) can be expressed as:

$$P_{\mathrm{x}} = \frac{1}{\sigma_{\mathrm{d,x}} + \mu_{\mathrm{d,x}}} \quad (5)$$

In this way it is possible to obtain high values of P for periodic and fast movements, that can be evidence of pathological motion patterns [13]. Finally, the periodicity of each spatial axes (x and y) and for both left (L) and right (R) limb are merged $P = \sum_s P_{\mathrm{s,L}} + \sum_s P_{\mathrm{s,R}}$, with $s = [x, y]$.

D. Motion patterns classification

Once we obtain the time sequences of all the points extracted from a video, we compute a feature vector of the 18 quantitative parameters: cross-correlation of hands speed, acceleration and jerk; cross-correlation of feet speed, acceleration and jerk; skewness of head, hands and feet speed; nose, hands and feet area out of standard deviation of moving average; nose, hands and feet area differing from moving average; nose, hands and feet periodicity.

Preliminary task: we first consider the classification between at-term and preterm infants as a sanity check for our pipeline. This experiment allows us to check if the devised procedure provides us with a reliable account of the infant's motion pattern, on a well defined and objective ground truth.

Main task: we then consider a medically-relevant task. We focus only on preterm infants and address the classification task of discriminating between the ones with and without neuro-motor disorders. In this case the ground truth was provided by physicians based on the Bayley test [4] performed two years after the video recording.

Classifier design: we compare different classifiers trained on the feature vector of parameters: a Random Forest (RF), a fully connected Neural Network (NN) with two hidden layers, and Support Vector Machines (SVM) with different kernels.

Greedy feature selection: we also reason on the discriminative potential of the features computed and the redundancy of the feature vector by applying a greedy feature selection procedure. The choice appears to be appropriate considering the small size of the feature vectors. We start by addressing a classification task based on a single (the best performing) parameter; then, we train a new classifier adding one parameter. At each iteration we add a parameter and we keep it in the list if the leave-one-out cross-validation accuracy of the updated classifier is higher than the one of the previous classifier.

III. Results

A. Preliminary Task: at-term vs preterm

We consider videos of 31 healthy infants: all the 13 born at-term and 18 born preterm randomly selected among the 28 with normal motion pattern. We first perform the classification task by considering all the extracted parameters with the different classifiers and leave-one-out cross-validation.

TABLE II
PRELIMINARY TASK. SUMMARY OF OVERALL ACCURACY (OA) FOR AT TERM / PRETERM CLASSIFICATION (PAR STANDS FOR PARAMETERS).

Classifier	*OA all par*	*OA best par*
Random Forest	**83.9**%	90.3%
Neural Network	77.4%	87.1%
SVM Polynomial Kernel	71.0%	**96.8%**
SVM Gaussian Kernel	67.7%	90.3%

We report the overall accuracy for all the tested classifiers on both the full feature vector and the best selected features in Table II. The reported results are very good and highlight a clear benefit in applying feature selection.

The classifier that allows to maximise the overall accuracy is the *SVM with polynomial kernel* with a sensitivity of 100% and a specificity of 94.4%. For this case the best parameters selected are: (1) cross-correlation of feet jerk, (2) hands speed skewness, (3) hands area out of standard deviation of moving average, (4) nose area differing from moving average, (5) hands area differing from moving average, (6) feet area differing from moving average, (7) nose periodicity and (8) hands periodicity.

B. Main Task: with vs without neuro-motor disorders

We consider 55 infants born preterm. We proceed as in preliminary task and we first consider all the motion parameters and then a subset of them. The results are reported in Table III.

TABLE III
MAIN TASK. SUMMARY OF OVERALL ACCURACY (OA) FOR THE CLASSIFICATION OF CHILDREN WITH AND WITHOUT NEURO-MOTOR DISORDERS (PAR STANDS FOR PARAMETERS).

Classifier	*OA all par*	*OA best par*
Random Forest	**56.4**%	69.1%
Neural Network	54.5%	74.5%
SVM Polynomial Kernel	50.9%	72.7%
SVM Gaussian Kernel	50.9%	**78.2%**

In this case, variable selection is even more crucial, since the results obtained by considering the full feature vector are only marginally above chance. Random forests are the only exception, since they naturally embed some form of variable selection. The classifier that maximizes the overall accuracy is the *SVM with gaussian kernel* and the subset of parameters selected are: (1) cross-correlation of hands speed, (2) cross-correlation of feet jerk, (3) hands speed skewness and (4) hands area differing from moving average. With this choice, we reach a sensitivity of 78.6% and a specificity of 77.8%.

C. Comparison with related works

In the reference field of this study, a comparison with other methods is not straightforward: several approaches differ significantly in spirit and on the acquisition devices adopted. Also, more importantly, for privacy reasons benchmark datasets are not available. Different studies evaluate different datasets, in terms of number of infants involved and in terms of level and intensity of neuro-motor disorders. We identify one method [13] which shares similarities with ours, in terms of parameters computed and classification tasks addressed, even if it is based on motion capture marker-based data (Vicon 370) and on a completely different infants population. The study involves a test-set composed by 14 infants (of which 11 healthy) acquired multiple times for a total of 52 acquisitions (46 of them related with healthy infants). The classification, also including a feature selection step and a quadratic discriminant analysis, lead to an overall accuracy on this test set of 73%. Instead, we consider 55 infants with and without neuro-motor disorders (one video recording for each infant) and we obtain the accuracy of 78.2%. These results suggest our marker-less approach is in line with the results obtained by more standard approaches, while reducing the obtrusiveness of the acquisition procedure.

IV. Discussion and Conclusions

In this paper we proposed a new automatic 2D maker-less pipeline to study infants' motion. Our pipeline does not require expensive and intrusive technologies, while reaching comparable performances. We decided to opt for a 2D pipeline in order to provide an easy to use system that -in our long term plan- could be adopted also by non-expert users. We tested the pipeline on a dataset of 68 children and we addressed two type of classification tasks. As a sanity check, we first discriminated between infants born at-term and preterm. This experiment allowed us to confirm the validity of our automatic analysis, based on quantitative motion parameters. We then evaluated the validity of the pipeline in the classification between infants with and without neuro-motor disorders: we relied on a clinically assessed ground truth provided by physicians after two years from birth as it is not easy to understand the neuro-motor status at 40 weeks of gestational age. We report encouraging results with an accuracy of 78.2%, speaking in favour of the method's potential as an early diagnosis tool.

In terms of future developments, one direction is to increase the set of parameters to include all the possible motion patterns described by the GMs theory (here we treat only some spontaneous movements [13]), this would also require a dense motion estimation [18]. We will also refine the classification layer by considering different levels of neuro-motor disorders. Finally, the research will take into account a longer term infants' follow-up, analysing videos acquired at different ages after birth to evaluate the evolution of the motion patterns.

References

[1] L Adde, J L Helbostad, A R Jensenius, G Taraldsen, and R Støen. Using computer-based video analysis in the study of fidgety movements. *Early human development*, 2009.

[2] M C Allen. Neurodevelopmental outcomes of preterm infants. *Current opinion in neurology*, 21(2):123–128, 2008.

[3] W Baccinelli, M Bulgheroni, V Simonetti, F Fulceri, A Caruso, L Gila, and M L Scattoni. Movidea: A software package for automatic video analysis of movements in infants at risk for neurodevelopmental disorders. *Brain Sciences*, 10(4):203, 2020.

[4] N Bayley. *Bayley scales of infant and toddler development: administration manual*. Harcourt assessment, 2006.

[5] Z Cao, T Simon, S-E Wei, and Y Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE CVPR*, pages 7291–7299, 2017.

[6] B Carse, B Meadows, R Bowers, and P Rowe. Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3d motion analysis system. *Physiotherapy*, 99(4):347–351, 2013.

[7] P Ciet, H AWM Tiddens, P A Wielopolski, J M Wild, E Y Lee, G Morana, and M H Lequin. Magnetic resonance imaging in children: common problems and possible solutions for lung and airways imaging. *Pediatric radiology*, 2015.

[8] S L Colyer, M Evans, D P Cosker, and A IT Salo. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open*, 4(1):24, 2018.

[9] N Hesse, C Bodensteiner, M Arens, U G Hofmann, R Weinberger, and A Sebastian Schroeder. Computer vision for medical infant motion analysis: State of the art and rgb-d data set. In *Proceedings of the ECCV*, pages 0–0, 2018.

[10] K Kawashima, Y Funabiki, S Ogawa, H Hayashi, Z Soh, A Furui, A Sato, T Shiwa, H Mori, K Shimatani, et al. Video-based evaluation of infant crawling toward quantitative assessment of motor development. *Scientific reports*, 2020.

[11] K W Krigger. Cerebral palsy: an overview. *American family physician*, 73(1):91–100, 2006.

[12] A Mathis, P Mamidanna, K M Cury, T Abe, V N Murthy, M W Mathis, and M Bethge. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, 21(9):1281, 2018.

[13] L Meinecke, N Breitbach-Faller, C Bartz, R Damen, G Rau, and C Disselhorst-Klug. Movement analysis in the early detection of newborns at risk for developing spasticity due to infantile cerebral palsy. *Human movement science*, 25(2):125–144, 2006.

[14] M Moro, G Marchesi, F Odone, and M Casadio. Markerless gait analysis in stroke survivors based on computer vision and deep learning: a pilot study. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 2097–2104, 2020.

[15] F B Palmer. Strategies for the early diagnosis of cerebral palsy. *The Journal of pediatrics*, 145(2):S8–S11, 2004.

[16] H F Prechtl. Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early human development*, 1990.

[17] DA Sival, GHA Visser, and HFR Prechtl. The effect of intrauterine growth retardation on the quality of general movements in the human fetus. *Early human development*, 28(2):119–132, 1992.

[18] A Stahl, C Schellewald, Ø Stavdahl, O M Aamo, L Adde, and H Kirkerod. An optical flow-based method to predict infantile cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(4):605–614, 2012.

[19] V A Stallings, C E Cronk, B S Zemel, and E B Charney. Body composition in children with spastic quadriplegic cerebral palsy. *The Journal of pediatrics*, 126(5):833–839, 1995.

[20] T Tsuji, S Nakashima, H Hayashi, Z Soh, A Furui, T Shibanoki, K Shima, and K Shimatani. Markerless measurement and evaluation of general movements in infants. *Scientific reports*, 10(1):1–13, 2020.