2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 979-8-3503-2447-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMBC40787.2023.10340961

Mean 3D Dispersion for Automatic General Movement Assessment of Preterm Infants

Ameur Soualmi[1], Olivier Alata[2], Christophe Ducottet[2], Hugues Patural[3], and Antoine Giraud[3]

***Abstract*— The General Movement assessment (GMA) is a validated assessment of brain maturation primarily based on the qualitative analysis of the complexity and the variation of spontaneous motor activity. The GMA can identify preterm infants presenting an early abnormal developmental trajectory before term-equivalent age, which permits a personalized early developmental intervention. However, GMA is time-consuming and relies on a qualitative analysis; these limitations restrict the implementation of GMA in clinical practice. In this study based on a validated dataset of 183 videos from 92 premature infants (54 males, 38 females) born <33 weeks of gestational age (GA) and acquired between 32 and 40 weeks of GA, we introduce the mean 3D dispersion (M3D) for objective quantification and classification of normal and abnormal GMA. Moreover, we have created a new 3D representation of skeleton joints which allows an objective comparison of spontaneous movements of infants of different ages and sizes. Preterm infants with normal versus abnormal GMA had a distinct M3D distribution (p <0.001). The M3D has shown a good classification performance for GMA (AUC=0.7723) and presented an accuracy of 74.1%, a sensitivity of 75.8%, and a specificity of 70.1% when using an M3D of 0.29 as a classification threshold.**

***Clinical relevance*— Our study paves the way for the development of quantitative analysis of GMA within the Neonatal Unit.**

I. INTRODUCTION

General movements represent spontaneous motor activity occurring from nine weeks of gestational age (GA) until the appearance of goal-directed movements around four months of corrected age [1]. The General Movement assessment (GMA) is a validated assessment of brain maturation primarily based on the qualitative analysis of the complexity and the variation of general movements [2]. The GMA of preterm infants performed prior to term and at term-equivalent age can identify those presenting an early abnormal developmental trajectory [3]. Early identification of such abnormal trajectory is critical to initiate a personalized early developmental intervention, and thus optimize the development of these children [4]. However, the implementation of GMA within the Neonatal Unit is limited because this assessment is time-consuming. Another limitation is that GMA is a qualitative assessment [1].

The recent progress in computer vision has led to many studies about GMA automation, from marker-based to markerless methods. Nelson et al. [5] have reviewed the different works in this area. Most of them focused on assessing the fidgety movements that appear from 2 months of age since it is difficult to get infants' videos before term age. Moreover, recent automatic GMA studies are oriented to 2D video analysis, using different kinds of both image features and estimated parameters [5]. The optical flow was widely used [6]–[8] to calculate some parameters such as the quantity of motion, centroid of motion, movement magnitude, etc. [9], [10]. 2D pose estimation was also investigated [11], [12] and others parameters like joints-angle [13] were introduced. Many have used deep learning algorithms to classify general movements as normal or abnormal [14], [15] but this kind of classification lacks interpretability. However, even though much effort was invested, 2D analysis remains limited since it does not exploit the overall infants' movement information in space, and information loss can go up to 53% due to dimensionality reduction [9]. Thus, many recent works have tried to use 3D automatic assessment with different methods. RGBD cameras are mostly used for 3D pose estimation, but it fails when dealing with occlusions and complex infant poses (see II-C). Wu et al. [16] introduced a complexity index but it lacks direct interpretation and was validated on a dataset of 12 infants only. More generally, the lack of annotated data is a common obstacle in this research area.

In this paper, we introduce the mean 3D dispersion as an objective quantification index for automatic GMA. This index is computed from the 3D pose of infants analyzed over a one-minute stereoscopic video. It relies on a new representation of the 3D skeleton joints over a unit sphere. Moreover, we propose an evaluation of this automatic assessment through a large dataset of 183 stereoscopic videos of 92 premature infants taken in a clinical environment and evaluated by an expert group composed of experienced General Movements Trust-certified assessors.

Our main contributions are: (i) to introduce the mean 3D dispersion (M3D) as a quantitative index of spontaneous motor activity of preterm infants, (ii) to propose a new normalized representation characterizing the activity of infants and based on the projection of the 3D skeleton joints over the unit sphere, (iii) to test the hypothesis that M3D can classify normal and abnormal GMA in an extensive and validated dataset of premature infants' videos acquired within the Neonatal Unit.

*A.S has benefited of a *Contrat doctoral de l'ED SIS.*

[1]Université Jean Monnet Saint-Étienne, INSERM, SAINBIOSE U1059, CNRS, Hubert Curien U5516, F-42023 Saint-Étienne, France ameur.soualmi@univ-st-etienne.fr

[2]Université Jean Monnet Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France olivier.alata@univ-st-etienne.fr; ducottet@univ-st-etienne.fr

[3]Université Jean Monnet Saint-Étienne, CHU Saint-Étienne, Service de Néonatalogie, INSERM, SAINBIOSE U1059, F-42023 Saint-Étienne, France hugues.patural@chu-st-etienne.fr; antoine.giraud@univ-st-etienne.fr

II. Methods

To objectively quantify the infants' spontaneous movements, we used an original method of 3D pose estimation using a stereoscopic framework (section II-C). Then, the 3D coordinates of the elbows, wrists, ankles, and knees were transformed to local origins on the infant's body to analyze their displacements according to the infant's position (sectionII-D.1). After that, since the size of the limbs is fixed, we used a unit sphere representation (sectionII-D.2) for each joint movement as a normalization procedure for our population. After filtering these directions on the sphere (section II-D.3) to get rid of duplications, we calculated the mean 3D dispersion parameter and performed our statistical analysis.

A. Ethics

This work is part of the AGMA study which was approved by an Ethical Committee in February 2021 (IDRCB 2020-A03335-34; *Comité de Protection des Personnes Sud-Est II*). Written parental consent was obtained from each participant.

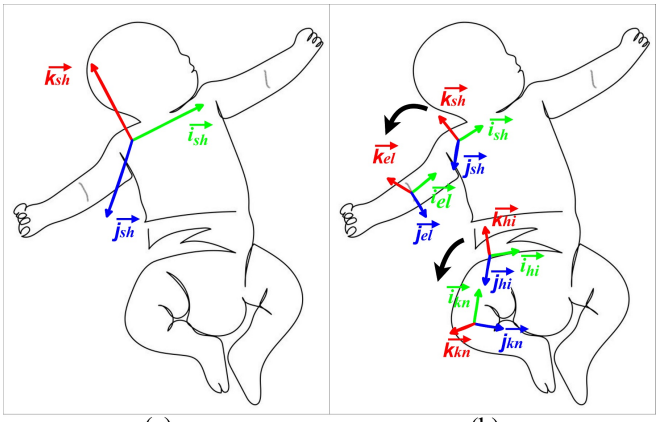
B. Dataset

Ninety-two preterm infants (54 males, 38 females) born <33 weeks of GA and hospitalized in the Neonatology department of the *Centre Hospitalier Universitaire de Saint-Étienne*, France, were included. A GMA was performed on 183 videos (132 with a normal GMA, 51 with an abnormal GMA) acquired between 32 and 40 (mean=36.3, SD=1.8) weeks of GA and classified by an expert group composed of experienced General Movements Trust-certified assessors, in accordance with the Prechtl's method of GMA [1].

C. 3D Pose Estimation

In order to assess the infants' movements in 3D, a particular framework was used which consists of recording videos of the babies with a stereoscopic camera (ZED2) at 30 FPS. Each pair of images was then used as input to a retrained convolutional neural network to get 2D pose estimations on each side view. The neural network used for this purpose was DARK [17] on top HRNet-W32 [18] which uses a new coordinate representation. This architecture was retrained on two common benchmark datasets (MPII and COCO) for the purpose of estimating the poses of adults only. We fine-tuned this model on an important dataset of real images of infants containing 4,250 stereoscopic images (88,500 in total) manually annotated and reviewed which improved its PCK@0.2 from 93.31% to 98.30%. Therefore, it became possible to automatically estimate 17 keypoints representing the 2D pose of infants without using any markers and with high precision. Hence, for each rectified stereo pair of infants' images, the corresponding 3D pose was obtained using triangulation on their respective 2D pose estimations.

Besides being safe and easy to implement, the particular advantage of this method over depth cameras is the ability to estimate the position of occluded joints in 3D since depth maps detect only visible body surfaces.



(a) (b)

Fig. 1. Origin Transformation. (a): Local origin for right elbow movement obtained using camera origin rotation and translation. (b): Transformation of shoulder and hip origins to elbow and knee origins.

D. 3D Representation

Analyzing the whole infant's movements in a video means taking into consideration the displacement of the upper and lower body members. Thus in our study, we specifically analyzed the movement of the elbows, wrists, knees, and ankles. For each stereoscopic video in our dataset, and applying our 3D pose estimation framework, the full 3D trajectories of the body joints in space were obtained. After applying a median filter with a kernel size of 3 to get rid of non-accurate points, we proceeded to analyze them as follows.

1) Origin Transformation: The 3D estimation process generates real-world coordinates of all keypoints in centimeters with respect to the left camera in the stereo-pair. Consequently, the same infant movements with different baby postures (lying on the back or side) can be considered different. To address this issue, we used a new local origin for each joint with respect to the infant's body.

- For the elbow movements, the shoulder on the respective side of the body was used as an origin. The vector $\vec{i_{sh}}$ connecting the two shoulders was used as the x-axis (abscissa). A cross-product between the two vectors linking the two shoulders with the point in the middle of the hips led to a second vector $\vec{k_{sh}}$ perpendicular to the infant's body plan and representing the z-axis (applicate). Lastly, the y-axis (ordinate) vector $\vec{j_{sh}}$ was obtained as the cross product of $\vec{i_{sh}}$ and vector $\vec{k_{sh}}$ after normalizing them. We refer to this new coordinate system by ($\vec{i_{sh}}$,$\vec{j_{sh}}$,$\vec{k_{sh}}$) for later usage (see Fig. 1).
- For the wrist movements, the elbow on the respective side of the body was used as an origin. The normalized vector $\vec{i_{el}}$ between the elbow and the shoulder on the same side was used as the x-axis. Thus, to get the other two vectors, we rotated and translated the previous coordinate system using Rodrigues' rotation formula. First, we started by finding the rotation axe $\vec{i_r}(x_r, y_r, z_r)$ as the cross product of $\vec{i_{sh}}$ and $\vec{i_{el}}$. Then

the rotation angle $\theta = \arccos\left(\vec{i_{sh}} \cdot \vec{i_{el}}\right)$ and finally the rotation matrix R calculated as follows: with $c = \cos\theta$, $s = \sin\theta$, and $c' = 1 - c$:

$$\mathbf{R} = \begin{bmatrix} c + x_r^2 c' & x_r y_r c' - z_r s & y_r s + x_r z_r c' \\ z_r s + x_r y_r c' & c + y_r^2 c' & y_r z_r c' - x_r s \\ -y_r s + x_r z_r c' & x_r s + y_r z_r c' & c + z_r^2 c' \end{bmatrix} \quad (1)$$

- For the knee movements, the hip on the respective side of the body was used as an origin. The vector $\vec{i_{hi}}$ connecting the two hips was used as the x-axis. A cross product between the two vectors linking the hips by the point in the middle of the shoulders led to a second vector $\vec{k_{hi}}$ perpendicular to the infant's body plan representing the z-axis. Lastly, the y-axis vector $\vec{j_{hi}}$ was obtained as the cross-product of the $\vec{i_{hi}}$ and $\vec{k_{hi}}$ vectors after normalizing them.
- For ankle movements, the knee on the respective side of the body was used as an origin. The normalized vector $\vec{i_{kn}}$ between the knee and the hip on the same side was used as the x-axis. Finally, to get the other two vectors, we rotated and translated the knee origin as we did for the shoulder origin.

2) Sphere Representation: For each joint, and after getting the new local coordinates throughout the whole duration of the video, we ended up with a point cloud representing the movements of the joint and consequently the limb with one end considered as the local origin. Therefore, the joint movement had a shape of a moving point on a sphere with a fixed radius equal to the limb length (see Fig. 2.a). The advantage of this representation is that it can easily be normalized to a unit sphere, eliminating the effect of varying height and body size between infants, and preserving the angles of the limb's movement. Therefore, an objective comparison of the quality of the general movements would be possible.

3) Angular Activation Map: The unit sphere obtained with the previous step represented all the joint positions in space from a one-minute duration video. It included the detections that resulted from very small movements which are not representative of GMs and produce duplicated positions. Hence, we converted the cartesian coordinates (x, y, z) of each point on the sphere to spherical coordinates (θ, φ) resulting in a distribution with two variables $0 \leq \theta < 2\pi$ and $0 \leq \varphi \leq \pi$. Then the support of this distribution was projected on an activation map with 128x64 dimensions ranging from $0 \rightarrow 2\pi$ and $0 \rightarrow \pi$ as shown in figure 2.c.

Therefore, for every point with spherical coordinates (θ, φ), the respective cell (u, v) in the activation map was set to 1 such that $\theta_{u-1} \leq \theta < \theta_u$ with $\left(\theta_u = \frac{\pi u}{64}, u = 1, 2, .., 128\right)$, and $\varphi_{v-1} \leq \varphi < \varphi_v$ with $\left(\varphi_v = \frac{\pi v}{64}, v = 1, 2, ..., 6\right)$, to eliminate the points duplication and noisy detections. Finally, the activated cells were converted back to cartesian coordinates resulting in filtered unit spheres (see Fig. 2.b) ready to be used for statistical analysis.

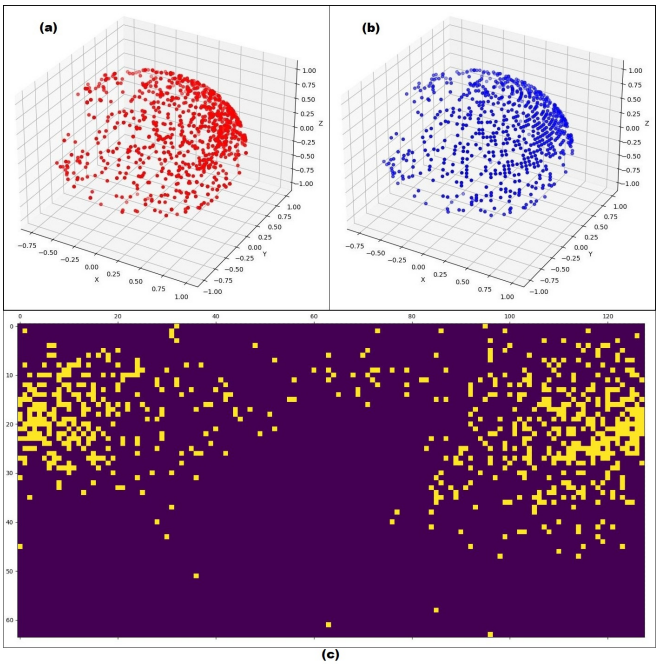


Fig. 2. (a): The unit sphere describing the movements of the infant's left hand around the elbow before filtering. (b): The unit sphere obtained after filtering. (c): The activation map resulted from sphere (a) points.

E. Mean 3D Dispersion Measure

To quantify how complex and multi-directional the infants' movements were in the space, the Mean 3D dispersion parameter was used. For all the 183 videos recorded, and for each joint, the 3D dispersion parameter was calculated using the filtered unit spheres representation described earlier, as follows: given a spherical point distribution with N points, with each point having an associated unit direction vector $\vec{v_i}$ (where $i = 1, ..., N$), the 3D polarization that describes the degree of alignment of points can be calculated as :

$$p = \frac{1}{N}\left|\sum_{i=1}^{N} \vec{v_i}\right| \quad (2)$$

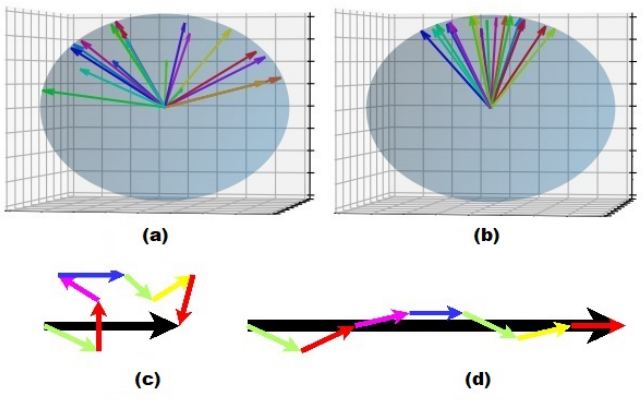


Fig. 3. (a) Dispersed points on a sphere. (b) polarized points on a sphere. (c) The sum of dispersed unit vectors. (d) The sum of polarized unit vectors.

Points that are concentrated in one particular direction as in Fig.3 will have an important polarization, and inversely the points that are well distributed all around the sphere will have low polarization. Hence, knowing that $0 < p < 1$, the 3D dispersion parameter $\sigma = 1 - p$ describes how well the points are scattered around the origin, which is the most suitable parameter for our study, denoting that a normal joint movement that is complex and goes in all directions will have a significant 3D dispersion compared to a movement that is repetitive and poor which is a key parameter in GMA.

F. Statistical Analysis

The two populations (normal and abnormal GMA) were analyzed and evaluated by calculating the 3D dispersion of each joint and then averaging them to get the M3D for every infant. The normalized histograms of averaged 3D dispersions distributions were calculated and then the kernel density estimate was plotted for each histogram (see Fig. 4). A Kolmogorov-Smirnov test was used to test the hypothesis of whether the dispersions from abnormal and normal GMs were from the same continuous distribution or not in addition to the asymptotic p-value. Also, the ROC with different thresholds and AUC were calculated. Statistical analysis was performed using Python 3.7. The significance level was set at 5%.

III. Results

Preterm infants with normal versus abnormal GMA had a distinct M3D distribution (p <0.001, see Fig.4). The classification performance analysis of M3D for GMA revealed an AUC of 0.7723. An accuracy of 74.1%, a sensitivity of 75.8%, and a specificity of 70.1% were obtained when using an M3D of 0.29 as a classification threshold (see Fig. 5) and considering the abnormal GMA as the positives and normal GMA as negatives.

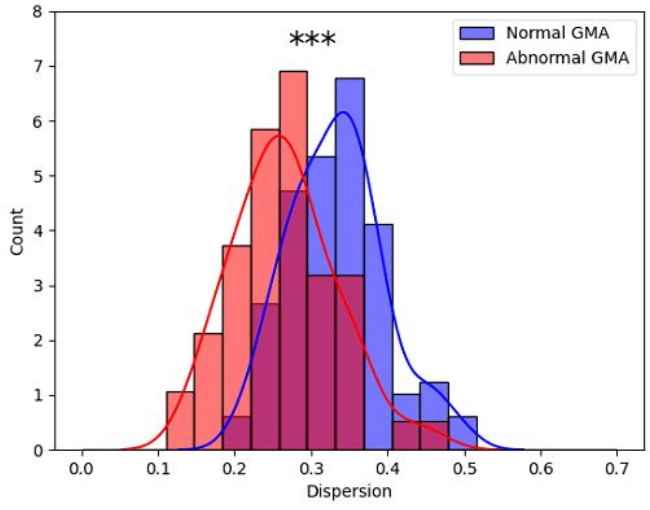


Fig. 4. Normalized histogram and kernel density plot of the mean 3D dispersion of preterm infants with normal (n=132) *versus* abnormal GMA (n=51). ***p $<$ 0.001, with a Kolmogorov-Smirnov Test. Abbreviation: GMA, General Movement assessment.

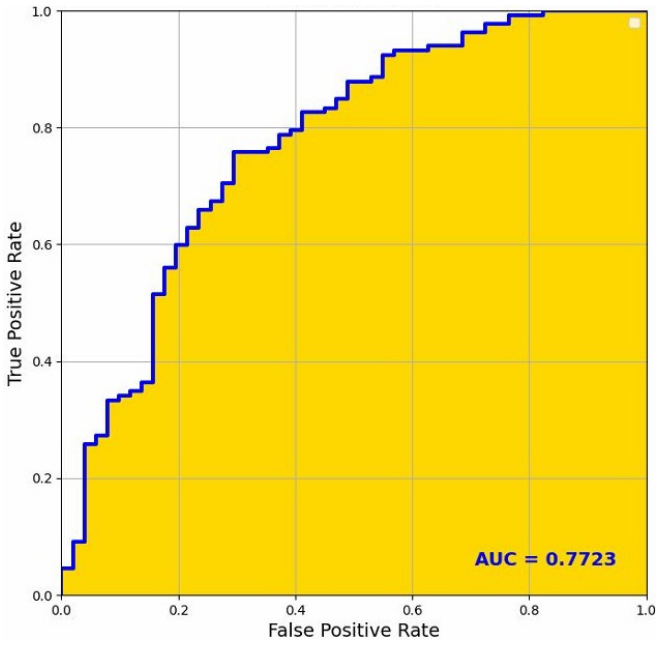


Fig. 5. ROC curve illustrating the classification performance of the M3D for GMA. Abbreviations: AUC, area under the ROC Curve; GMA, General Movement Assessment; M3D, mean 3D dispersion.

IV. Discussion

The M3D displayed a good classification performance for the GMA of preterm children. With an AUC of 0.77 (see Fig 5), a threshold of 0.29 was chosen as the best classification index with the highest accuracy reported on such an important dataset. Previous studies used the MINI-RGBD dataset, which contains 12 videos of 1000 frames ($\approx$ 33s) each, representing synthetic preterm infants with real movements. The study on 3D GMA classification from Wu et al. [16] introduced an evaluation index S calculated using a single angle per join. McCay et al. [14] presented neural network architectures for classifying pose-based features as normal and abnormal GMA. They provided both accurate results but it was not possible to make an objective comparison due to the type and the small number of videos used. Moreover, these two studies used the Openpose network [19] which was trained on human adult images only, therefore raising the risk of keypoints detection error. Whereas for our study, in addition to the fact that it is a 3D analysis method, our pose estimation model was retrained on a manually annotated dataset of 88k real infant images. Moreover, the 3D mean dispersion parameter was validated on a dataset of 183 videos of infants with homogeneous gestational age and classified by an expert group composed of experienced General Movements Trust-certified assessors. Also, the spherical representation used had a major impact on the ability to compare these populations with different weights and ages since we normalized on the actual size of the infants' limbs. Hence, we were able to avoid biased quantification for infants with larger limbs and consequently larger movement amplitudes. This allowed for a more accurate and reliable comparison of the two populations. Moreover, its usage can

be extended to methods where the absolute distance between joints cannot be measured, such as pixels that depend on several factors like camera resolution and camera–subject distance, making the measurements not constant outside the single video framework and the comparison of data among different subjects not valid [9]. Yet one limitation of this study is that the 3D dispersion parameter can describe the spatial characteristics and the complexity of infant movements but not their fluidity and variability which are important parameters for GMA. The perspective of this work is to develop other parameters which can describe the fluidity and the variability of the general movements by analyzing the temporal characteristics such as velocities and acceleration vectors. These other parameters would allow a better classification of normal and abnormal general movements.

ACKNOWLEDGMENT

We are grateful to all the members of the Saint-Étienne GMA team.

REFERENCES

[1] G. Cioni, F. Ferrari, A. F. Bos, H. F. R. Prechtl, and C. Einspieler, *Prechtls Method on the Qualitative Assessment of General Movements in Preterm, Term and Young Infants*. Mac Keith Press, 2008.

[2] M. Hadders-Algra, "Neural substrate and clinical significance of general movements: an update," *Developmental Medicine & Child Neurology*, vol. 60, 2018.

[3] J. E. Olsen, J. L. Y. Cheong, A. L. Eeles, T. L. Fitzgerald, K. L. Cameron, R. Albesher, P. J. Anderson, L. W. Doyle, and A. J. Spittle, "Early general movements are associated with developmental outcomes at 4.5-5 years." *Early human development*, vol. 148, p. 105115, 2020.

[4] A. Spittle, J. Orton, P. Anderson, R. Boyd, and L. Doyle, "Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants," *The Cochrane Library*, vol. 2015, 2015.

[5] N. Silva, D. Zhang, T. Kulvicius, A. Gail, C. Barreiros, S. Lindstaedt, M. Kraft, S. Bölte, L. Poustka, K. Nielsen-Saines, F. Wörgötter, C. Einspieler, and P. Marschik, "The future of general movement assessment: The role of computer vision and machine learning – a scoping review," *Research in Developmental Disabilities*, vol. 110, p. 103854, 2021.

[6] S. Orlandi, K. Raghuram, C. R. Smith, D. Mansueto, P. Church, V. Shah, M. Luther, and T. Chau, "Detection of atypical and typical infant movements using computer-based video analysis," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 3598–3601.

[7] K. Raghuram, S. Orlandi, V. Shah, T. Chau, M. Luther, R. Banihani, and P. Church, "Automated movement analysis to predict motor impairment in preterm infants: a retrospective study," *Journal of Perinatology*, vol. 39, pp. 1–8, 2019.

[8] E. A. F. Ihlen, R. Støen, L. Boswell, R.-A. de Regnier, T. Fjørtoft, D. J. Gaebler-Spira, C. Labori, M. Loennecken, M. E. Msall, U. I. Möinichen, C. Peyton, M. Schreiber, I. E. Silberg, N. T. Songstad, R. T. Vågen, G. K. Øberg, and L. Adde, "Machine learning of infant spontaneous movements for the early prediction of cerebral palsy: A multi-site cohort study," *Journal of Clinical Medicine*, vol. 9, 2019.

[9] 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*, vol. 10, no. 4, 2020.

[10] 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*, vol. 10, no. 1, 2020.

[11] K. D. McCay, P. Hu, H. P. H. Shum, W. L. Woo, C. Marcroft, N. D. Embleton, A. Munteanu, and E. S. L. Ho, "A pose-based feature fusion and classification framework for the early prediction of cerebral palsy in infants," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 8–19, 2022.

[12] D. Sakkos, K. McCay, C. Marcroft, N. Embleton, S. Chattopadhyay, and E. Ho, "Identification of abnormal movements in infants: A deep neural network for body part-based prediction of cerebral palsy," *IEEE Access*, vol. 9, pp. 94 281–94 292, 2021.

[13] H. I. Shin, H.-I. Shin, M. S. Bang, D.-K. Kim, S. H. Shin, E.-K. Kim, Y.-J. Kim, E. S. Lee, S. G. Park, H. M. Ji, and W. H. Lee, "Deep learning-based quantitative analyses of spontaneous movements and their association with early neurological development in preterm infants," *Scientific Reports*, vol. 12, no. 1, p. 3138, 2022.

[14] K. D. McCay, E. S. L. Ho, H. P. H. Shum, G. Fehringer, C. Marcroft, and N. D. Embleton, "Abnormal infant movements classification with deep learning on pose-based features," *IEEE Access*, vol. 8, pp. 51 582–51 592, 2020.

[15] S. Reich, D. Zhang, T. Kulvicius, S. Bölte, K. Nielsen-Saines, F. Pokorny, R. Peharz, L. Poustka, F. Wörgötter, C. Einspieler, and P. Marschik, "Novel ai driven approach to classify infant motor functions," *Scientific Reports*, vol. 11, no. 1, p. 9888, 2021.

[16] Q. Wu, G. Xu, F. Wei, C. Longting, and S. Zhang, "Rgb-d videos-based early prediction of infant cerebral palsy via general movements complexity," *IEEE Access*, vol. 9, pp. 42 314–42 324, 2021.

[17] F. Zhang, X. Zhu, H. Dai, M. Ye, and C. Zhu, "Distribution-aware coordinate representation for human pose estimation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 7091–7100.

[18] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 5686–5696.

[19] Z. Cao, G. Martinez, T. Simon, S.-E. Wei, and Y. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2021.