



# HHS Public Access

Author manuscript

*Res Dev Disabil.* Author manuscript; available in PMC 2021 March 24.

Published in final edited form as:

*Res Dev Disabil.* 2021 March ; 110: 103854. doi:10.1016/j.ridd.2021.103854.

## The future of General Movement Assessment: The role of computer vision and machine learning – A scoping review

Nelson Silva<sup>a,b,1</sup>, Dajie Zhang<sup>a,c,d,1</sup>, Tomas Kulvicius<sup>e</sup>, Alexander Gail<sup>d,f</sup>, Carla Barreiros<sup>b,g</sup>, Stefanie Lindstaedt<sup>b,g</sup>, Marc Kraft<sup>h</sup>, Sven Bölte<sup>i,j,k</sup>, Luise Poustka<sup>c,d</sup>, Karin Nielsen-Saines<sup>l</sup>, Florentin Wörgötter<sup>d,e,m</sup>, Christa Einspieler<sup>a</sup>, Peter B. Marschik<sup>a,c,d,i,\*</sup>

<sup>a</sup>iDN – Interdisciplinary Developmental Neuroscience, Division of Phoniatics, Medical University of Graz, Graz, Austria

<sup>b</sup>Know-Center GmbH, Graz, Austria

<sup>c</sup>Child and Adolescent Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany

<sup>d</sup>Leibniz-ScienceCampus Primate Cognition, Göttingen, Germany

<sup>e</sup>Department for Computational Neuroscience, Third Institute of Physics-Biophysics, Georg-August-University of Göttingen, Göttingen, Germany

<sup>f</sup>German Primate Center – Leibniz Institute for Primate Research, Göttingen, Germany

<sup>g</sup>Institute of Interactive Systems and Data Science, Graz University of Technology, Graz, Austria

<sup>h</sup>Department of Medical Engineering, Technical University Berlin, Berlin, Germany

<sup>i</sup>Center of Neurodevelopmental Disorders (KIND), Centre for Psychiatry Research, Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

<sup>j</sup>Child and Adolescent Psychiatry, Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden

<sup>k</sup>Curtin Autism Research Group, School of Occupational Therapy, Social Work and Speech Pathology, Curtin University, Perth, Western Australia, Australia

<sup>l</sup>Division of Pediatric Infectious Diseases, David Geffen UCLA School of Medicine, USA

<sup>m</sup>Institute of Physics, Department for Computational Neuroscience at the Bernstein Center Göttingen, Georg-August-University of Göttingen, Göttingen, Germany

### Abstract

**Background:** The clinical and scientific value of Prechtl general movement assessment (GMA) has been increasingly recognised, which has extended beyond the detection of cerebral palsy

---

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

\*Corresponding author at: Child and Adolescent Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany. [peter.marschik@med.uni-goettingen.de](mailto:peter.marschik@med.uni-goettingen.de) (P.B. Marschik).

<sup>1</sup>These authors contributed equally.

Declaration of Competing Interest

We declare that there are no conflicts of interest, guiding this research work.

throughout the years. With advancing computer science, a surging interest in developing automated GMA emerges.

**Aims:** In this scoping review, we focused on video-based approaches, since it remains authentic to the non-intrusive principle of the classic GMA. Specifically, we aimed to provide an overview of recent video-based approaches targeting GMs; identify their techniques for movement detection and classification; examine if the technological solutions conform to the fundamental concepts of GMA; and discuss the challenges of developing automated GMA.

**Methods and procedures:** We performed a systematic search for computer vision-based studies on GMs.

**Outcomes and results:** We identified 40 peer-reviewed articles, most (n = 30) were published between 2017 and 2020. A wide variety of sensing, tracking, detection, and classification tools for computer vision-based GMA were found. Only a small portion of these studies applied deep-learning approaches. A comprehensive comparison between data acquisition and sensing setups across the reviewed studies, highlighting limitations and advantages of each modality in performing automated GMA is provided.

**Conclusions and implications:** A “method-of-choice” for automated GMA does not exist. Besides creating large datasets, understanding the fundamental concepts and prerequisites of GMA is necessary for developing automated solutions. Future research shall look beyond the narrow field of detecting cerebral palsy and open up to the full potential of applying GMA to enable an even broader application.

## Keywords

Augmented general movement assessment; Automation; Cerebral palsy; Computer vision; Deep learning; Developmental disorder; Early detection; General movements; Infancy; Machine learning; Neurodevelopment; Pose estimation

## 1. Introduction

Early detection of developmental disorders of various aetiologies, which are usually diagnosed during toddler-years or older, is a major challenge to clinicians and scientists across disciplines. Over the years, this field has become increasingly complex and has incorporated developmental, clinical, as well as technical perspectives. Besides the classic biomarker approaches targeting earlier identification of such *late detected developmental disorders* (LDDDs), the assessment of neurofunctional or behavioural biomarkers has caught increasing attention (e.g., Varcin & Nelson, 2016; Marschik, Einspieler, Sigafos, Enzinger, & Bölte, 2016; Marschik et al., 2017; Peyton & Einspieler, 2018). Research in different behavioural domains from early life and onwards has adopted both retrospective and prospective paradigms, such as the retrospective work on Rett syndrome (e.g., Einspieler & Marschik, 2019), or the ever-growing field of prospective siblings studies on autism spectrum disorders (e.g., Ali, Charman, Johnson, Jones, & Team, 2020; Bölte et al., 2016; McDonald et al., 2020; Murphy & Sporeen, 2012; Ozonoff et al., 2015; Shephard et al., 2019)

In this scoping review, we address one specific behavioural domain, the developing motor functions in the first few months of life. We focus on the subdomain of spontaneous *general movements* (GMs) and aim to recapitulate current computer vision-based studies on tracking and detection of GMs.

First operationalised by Heinz Prechtl and colleagues (e.g., Einspieler, Marschik, & Prechtl, 2008; Prechtl, 1990; Prechtl et al., 1997), the assessment of GMs has opened a unique window for scientists and clinicians to sight with their bare eyes the integrity of the young developing nervous system. Our interdisciplinary developmental neuroscience lab and the systemic ethology and development research lab, originated and founded by Heinz Prechtl and Christa Einspieler, inherit the long tradition and rich experience of studying GMs and bear the mission to extend the knowledge of GMs. Maintaining the high standard of the Prechtl general movement assessment (GMA), it is our vision to translate the classic GM field, the prediction of cerebral palsy (CP), to broader applications, incorporating innovative routes and wider perspectives.

GMs are but a part of the spontaneous movement repertoire (i.e., not induced by any external stimulus) and are present from early foetal life towards the end of the first half-year postterm. GMs involve the entire body, hence the term *general* movements. GMs are variable sequences of movements of the arm, leg, neck, and trunk with changing intensity, force, and speed (e.g., Einspieler, Marschik et al., 2008; Einspieler & Prechtl, 2005). A sequence of GMs waxes and wanes gradually, involving fluent and elegant rotations along the limbs' axis and slight changes in the movement direction. GMs are complex in appearance, and importantly, variable. When the developing nervous system is impaired, GMs lose complexity; their smooth and variable character alters and becomes monotonous, abrupt, or disorganised. Importantly, GMs present distinct age-specific patterns during the pre-term and term periods, and at 3–5 months of age. While at term age and shortly after, the writhing movements (WMs) dominate, the fidgety movements (FMs) gradually set in between 6–8 weeks, become pronounced at 12–16 weeks, and vanish around 20 weeks of postterm age (PTA). The quality of GMs can be examined by the Prechtl GMA, one of the most sensitive and reliable diagnostic tools for the prediction of cerebral palsy (e.g., Kwong, Fitzgerald, Doyle, Cheong, & Spittle, 2018; Novak et al., 2017; Prechtl et al., 1997). Quality of GMs is defined into age-specific normal vs abnormal categories. Abnormal GM patterns during the writhing movement period include: *poor repertoire*, *cramped-synchronized*, or *chaotic* movements; and during the fidgety movement period: *abnormal* or *absent* fidgety movement patterns. Especially, normal FMs suggest normal neurological development while the absence of FMs at 3–5 months PTA is the most sensitive and specific indicator of later neurological impairments, such as cerebral palsy (e.g., Bosanquet, Copeland, Ware, & Boyd, 2013; Einspieler, Bos et al., 2019; Einspieler, Utsch et al., 2019; Einspieler, Peharz, & Marschik, 2016; Einspieler & Prechtl, 2005; Kwong et al., 2018; Prechtl et al., 1997).

Initially a powerful predictor of CP, general movements have been studied worldwide in a multitude of neurodevelopmental and genetic disorders (e.g., Herrero et al., 2017; Romeo et al., 2008; Tomantschger et al., 2018). Accumulating evidence reveals elevated occurrences of aberrant GMs in infants later diagnosed with LDDDs, e.g., autism spectrum disorder, Rett syndrome (e.g., Einspieler et al., 2014; Einspieler & Marschik, 2019; Zappella et al., 2015),

or a range of early-identifiable disorders such as Down syndrome (e.g., Herrero et al., 2017) and Cornelia de Lange syndrome (e.g., Marschik, Soloveichick, Windpassinger, & Einspieler, 2015). Abnormal GMs are also present in infants born to mothers with viral infections like HIV or Zika that affect the central nervous system (e.g., Brasil et al., 2016; Einspieler, Utsch et al., 2019; Einspieler & Marschik, 2020a, 2020b; Palchik, Einspieler, Evstafeyeva, Talisa, & Marschik, 2013; Soares-Marangoni et al., 2019). The significance of GMs in early brain development in general and, consequentially, its long-term relevance for the later development of cognitive, speech-language, and motor functions has been increasingly recognised (e.g., Einspieler, Bos, Libertus, & Marschik, 2016; Einspieler, Peharz et al., 2016; Grunewaldt et al., 2014; Salavati et al., 2017). Although abnormal GMs, especially the absence of FMs during 3–5 months, do not point to a specific disorder, they flag high risks for future neurological impairments. If GMA could be manualised in daily clinical routines, it would support the earlier identification of LDDD and other neurodevelopmental impairments. Infants identified with abnormal GMs would be monitored more closely, and could thus, be referred sooner for specific diagnostic evaluations and benefit earlier from interventions (e.g., Peyton & Einspieler, 2018; Zang et al., 2016).

As GMA requires only 3–5 minutes observation of an infant's spontaneous movement (i.e., the infant needs not to be touched by the assessor), it is an evaluation far easier to be carried out than most assessments for neurological development. Hence GMA is suitable for daily clinical applications, particularly in low-resource settings. Being entirely non-intrusive, GMA is widely accepted by caregivers with divergent social and cultural backgrounds (e.g., Burger & Louw, 2009; Soleimani, Teymouri, & Biglarian, 2013; Tomantschger et al., 2018).

However, GMA can only be performed by certified assessors. Acquiring specific high-quality training is a prerequisite for a GMA assessor, and regular practices and recalibrations are indispensable. This is one reason why GMA has not yet been established universally in the daily clinical routines. Although interrater reliability of GMA has proven to be excellent across various studies at different sites (e.g., Einspieler & Prechtel, 2005; Kwong et al., 2018; Valle, Støen, Sæther, Jensenius, & Adde, 2015; Yuge et al., 2011), assessor skills surely vary from individual to individual and can be influenced by adverse human or environmental factors. So much as the clinical and scientific credit of GMA has been acknowledged, we need complementing avenues to scale up this valuable tool, where modern technology may be able to play a more decisive role. Indeed, in the past two decades, a boom of technological approaches aiming at automated or technology-assisted GMA have surfaced. These efforts range from mobile-app-based recording tools, e.g., the Baby Moves (Spittle et al., 2016) and the GMAApp (Marschik, Pokorny, et al., 2017), to automated pose estimation through sensor-based or markerless approaches (e.g., Irshad, Nisar, Gouverneur, Rapp, & Grzegorzec, 2020; Marcroft, Khan, Embleton, Trenell, & Plötz, 2015; Marschik et al., 2017).

In this paper, we provide an in-depth analysis of the most recent technology-driven studies on GMs. We focus on video-based approaches only, since GMA is in origin a visual-based method. Advanced computer vision technology remains authentic to the non-intrusive character of the classic GMA, allowing automated analyses of the infant's spontaneous

movements, which is not influenced by the use of wearable sensors and other devices (for recent overviews on diverse sensors targeting GMs, please see Hyde et al., 2019; Irshad et al., 2020; Marcroft et al., 2015). Different from previous reviews, we examine in particular if the existent technological attempts targeting automated GMA are in accordance with the fundamental concepts underlying this unique clinical tool. Specifically, we aim to: (1) provide an overview of available video-based approaches targeting GMs; (2) identify their techniques for movement detection, tracking, data pre-processing, and classification; and most importantly, (3) discuss from both, the conceptual and the technological perspectives, the major challenges, as well as advantages of incorporating automated visual-based approaches into classic GMA to enable an even broader application in daily clinical routines.

## 2. Materials and methods

### 2.1. Search methods

A search with thirteen well-known databases and research networks (Fig. 1) was carried out in September 2020. Fig. 1 summarises the complete search and screening procedure. In addition to the thirteen different sources, we also searched in Google, including personal webpages, blogs, forums, thesis, patents, and performed ancestral research of published papers to collect additional studies.

Following our aims, we defined three core categories of interests (COIs) for the search process: general movements, machine learning, and computer vision. Search terms of each COI are presented in Table 1. Studies published in English and found to be related to the three COIs were all collected and organised first via Zotero (Zotero (about), 2020). All the records were subsequently exported from Zotero to the visualisation tool SurVis (Beck, Koch, & Weiskopf, 2016) for automatic analysis of publication dates, keywords, authors, and topic clusters. We applied text analysis (using R) to examine the full texts (e.g., leading journals, top keywords). Our search concentrated on technological approaches and studies of infants, covering applications on both the automated analysis of movements and the early detection/prediction of developmental disorders. The search resulted in a total of 433 relevant records. In the following step, we screened these 433 records.

### 2.2. Screening

First, all duplicates and non-peer-reviewed articles were deleted. Second, articles published before 2010 were excluded to focus on the significant technological advancements during the past decade. Third, we removed studies of older infants (participants were on average 6 months of age or older); studies not targeting GMs; studies that did not apply machine learning; or studies that did not use video-based techniques.

## 3. Results

According to our search and screening procedures, we identified 40 peer-reviewed articles, 10 being conference contributions. All of these studies provided in-depth technical and algorithmic details on infant movement analysis and applied automated video-based approaches with machine and deep learning techniques (Table 2). Most studies ( $n = 30$ ) were

published between 2017 and 2020, indicating a strong interest in, and a boom of, video-based approaches addressing GMA in the recent years.

Table 2 provides a detailed overview of the various studies targeting automated GMA, summarising their approaches and techniques on infant movement detection, tracking, and classification. Table 3 provides a comprehensive comparison between data acquisition and sensing setups across the reviewed studies, highlighting limitations and advantages of each modality in performing automated GMA.

In Table 2, we split the studies into two generic groups: conventional machine learning models (CML,  $n = 35$ ; e.g., SVM, random forest) and deep learning models (DL,  $n = 5$ ; e.g., CNN, LSTM). In regards to tracking and detection techniques, 12 CML studies and 4 DL studies applied pose estimation (with OpenPose or a custom pose implementation). The other 24 CML studies applied diverse tracking and detection methods.

Multiple motion-related techniques were exploited, such as Optical Flow, i.e., a technique for tracking the motion of an infant across multiple frames to estimate the velocity of body parts and predict the position of each body part in the next frame, or a Particle Filter used as a technique for localisation and mapping in Optical Flow, or Graph-cut, a graph-based segmentation technique used before executing a Particle Filter.

As presented in Table 2, a variety of movement features were extracted, such as kinematic features (i.e., standard or customised features that define velocity and acceleration of points in a moving body); frequencies, amplitudes, and covariation of movement parameters (e.g., position, velocity, or acceleration); other spectral components (e.g., harmonics in periodic vibrations in moving body parts, used for FMs detection). Using pose estimation, Orlandi et al. (2018) developed a new set of time-related features to detect FMs. Moccia, Migliorelli, Carnielli, and Frontoni (2020), with a newly invented “*Pose Tool*”, calculated the standard deviation of joint angles over time by using visual indicators to represent such deviations. Cenci, Liciotti, Frontoni, Zingaretti, and Carnielli (2017) introduced a new movement state-vector to their model defining whether a targeted body part *is* or *is not* in motion by modelling the infant’s movement sequence as a series of transitional states using a Markov Chain (MC).

To categorise these features, diverse computational algorithms were used, such as KAZE, i.e., a multiscale 2D feature detection and description algorithm (Alcantarilla, Bartoli, & Davison, 2012), Large Displacement Optical Flow (LDOF), i.e., an integration of rich descriptors into a variational optical flow setting to detect small-fast moving body parts (J. M. Brox, 2011; T. Brox, Bruhn, Papenbergh, & Weickert, 2004), Markov Random Fields (MRF), i.e., used to encode contextual constraints into the prior probability (Pal & Pal, 1993), and Random Spectral Regression (RSR), i.e., a human action recognition algorithm based on random spectral regression (Lin, Zhu, Fan, & Fan, 2011).

#### 4. Discussion

Over the past decade, the significant clinical and scientific value of the Prechtl GMA has been increasingly recognised. Armed with the rapid advancing computer science, a surging

interest in developing automated GMA prevails in the field. Among the identified studies directly devoted to automated vision-based GMA, the majority were published within the past five years, and more are coming day after day (e.g., Doroniewicz et al., 2020; Groos, Adde, Støen, Ramampiaro, & Ihlen, 2020). As a limitation, we targeted only the publications in English during the past decade. Some work in the field could hence have escaped our review. Still, this scoping review, which aims at mapping the key concepts underpinning the research area of vision-based GMA, reflects on the cutting-edge of the field. In this section, we discuss the current approaches addressing automated solutions of GMA from both conceptual and technological perspectives.

#### 4.1. Conceptual considerations

Given the expanding interests in automated movement analysis, any attempt to develop computer-driven GMA requires a genuine understanding of the underlying concepts of the GMs and an intensified scientific sensitivity. GMA is by nature *gestalt*, the perception and interpretation of the infant's entire movement pattern without emphasising isolated parts. By contrast, computer-based methods are built upon minute *features* to generate algorithms. Although the automated GMA aims at overall classification, it remains a critical question, if and how human gestalt perception can be appropriately emulated by artificial intelligence (AI)? To validate tech-driven GMA, not only the interpretation of the classes, but more importantly, the extracted features, especially those obtained with unsupervised machine learning techniques, are of great conceptual, theoretical, and clinical importance. Otherwise, we might end up with merely a handful of discrete labels while losing the essential scientific and clinical semantics of GMA. To this end, we would need a more open communication between GMA experts and computer scientists to ensure the validity of future computerised models.

Speaking of the fundamental concepts of GMA, GMs are a significant constituent of the young infants' broad spontaneous movement repertoire and must be observed within the specific age span. As introduced at the beginning, infant movement patterns change dramatically during the very first months of life. Movements around term age are qualitatively different from the ones during the 3–5 month period, as these motor patterns mirror the developmental status of the nervous system at each respective age. Unfortunately, essential information on the participants' characteristics (e.g., the gestational age) was frequently missing in the discussed studies (Table 2). Some studies, although technically related to automated GMA, sampled infants beyond the age at which the GMs could be observed (e.g., Ouss et al., 2018, 2020). This implies that the classification algorithms of these studies might have been built upon (at least partly) inappropriate inputs, and the prediction would then have little to do with GMA per se. Relatedly, the current motion-tracking libraries and frameworks are mostly based on models for tracking adult movements, which are inherently different from those of the young infants. There is a need for further exploration as to how and if these "large-body oriented" motion tracking frameworks could be adapted to track infants' body parts and joints, as well as their motor specificities with suitable recording setups. *Infant-specific* models are needed in their own right to account for the subject's age-specific anatomical and motor constraints (Hesse, Pujades et al., 2018; Hesse, Schroeder et al., 2018; Ihlen et al., 2020).

Needless to say, computational models can only make predictions based on the datasets they are trained on, no more and no less. The nature of the input for training a model inevitably determines the validity and quality of its output. While attempting to acquire data for creating algorithms for automated GMA, we face the following challenges:

**4.1.1. Sample attributes**—Besides the age-specificity issue discussed above, we need to ask which high-risk groups or disorders are targeted (e.g., preterm-infants, who are at elevated risk for developing CP)? Is an adequate and appropriate control group (i.e., typically developing infants) included, which is important for all machine learning methods (Schmidt, Regan, Fahey, & Paplinski, 2019)? Is the sample representative for the targeted group and the sample size (number of participants and the amount of data from each infant) sufficient, so that the outcome is reliable and generalisable? We should not forget that GMs have a large, complex, and variable repertoire, bringing difficulties for machine learning approaches to acquire a representative dataset. For example, when relying on retrospective videos from infants with atypical development, due to the uncertain representativeness of the training datasets, it might be challenging to achieve high external validity when testing the created model on novel samples (Irshad et al., 2020).

**4.1.2. Sensing and recording setups**—As previously mentioned, despite the type of camera setups (Table 3), the non-intrusive classic GMA requires a standard viewing perspective to observe the infant's entire body. The infant is in supine position and untouched, moving free of any external stimulus, and should also be in an appropriate behavioural state (Einspieler et al., 2014). Otherwise, the movement pattern could be distorted. To maintain the non-intrusive character of the GMA, vision-based markerless approaches appear more favourable than the ones using wearable sensors, or attaching markers to the infant's body. Although marker- and wearable sensor-based approaches have technical merits (see Irshad et al., 2020), it is yet to be examined whether these markers or sensors may interfere with the infants' spontaneous movements, or whether the device-attaching procedures, usually time-consuming and during which the infant has to be touched or manipulated, could affect the infant's consequential behavioural state (e.g., becoming fussy and distracted).

**4.1.3. Dataset annotation and segmentation**—The quality of the annotation, being a key for the machine learning training dataset and the basis for classification, is largely neglected in the majority of the reviewed articles. In most cases, no information was provided on whether the dataset was annotated by certified GMA assessors, let alone the inter- and intra-rater reliability of the annotation by the GMA assessors. At the moment, no expert-annotated and validated public accessible large datasets are yet available for the purpose of scientific research. To realise automated GMA, creating such datasets might be challenging, partly due to complex confidentiality and privacy issues, which are however indispensable.

During the data annotation procedure, the duration of the video segments to be labelled is another puzzling issue. For machine learning methods, the shorter the movement duration, the less complex the model (i.e., less parameters), and thus, the shorter the time needed for training the model (assuming that shorter movement durations lead to smaller feature



vectors). For human assessors, however, a 2–5 min observation is normally required by the GMA standard to evaluate the infant’s movement repertoire. It is yet to be explored, if human assessors are capable of annotating very short video clips confidently and reliably. More importantly, it is normal if a desired type of movements (targeted by the computer model), for example, the FMs, is absent for a short interval (e.g., 5 s) in a fully typically developing infant, who is in the fidgety movement age period. If the data annotation would be based, for example, on 20-second clips each, where both the fidgety movements (“1’’) and non-fidgety movements (“0’’) could occur, an annotation of either “0’’ or “1’’ for the respective 20 s could be inappropriate. That said, we need to find a compromise between a reasonable unit duration, appropriate feature encoding for machine learning algorithm and a minimum length of video for human assessors to be able to evaluate.

#### 4.2. Technological considerations

From a technological perspective, a wide variety of sensing, tracking, detection, and classification tools for automated GMA based on computer vision are available (Table 2). Not only research approaches are heterogeneous, their datasets for training and testing across studies are also divergent. For this reason, a cross-study comparison on the model performance is almost impossible. Only a small portion of the existing studies applied deep learning approaches (DL,  $n = 5$ ), which is likely to change in the near future. DL, being able to extract latent data features in an unsupervised way (e.g., using autoencoder architectures), is more suitable for handling massive datasets to achieve high performance. Efforts on creating larger validated datasets are needed and will allow further advancements in developing the DL models.

Given the various techniques applied, no current automated solution could yet defeat human GMA experts. Consequently, a fully automated GMA for the clinical practice seems rather elusive in the near future. To increase the performance of the technical approaches, on one hand, we need to better comprehend the underlying principles of the classic GMA, create larger annotated valid datasets, and revisit the capability and limitation of the existing approaches; on the other hand, we might need to develop novel strategies. For example, in addition to traditional methods to prevent overfitting such as training using early validation stop or utility of drop-out layers in DL, we could introduce additional regularisation methods (e.g., noise injection; Kuka ka, Golkov, & Cremers, 2017) to the models to reduce overfitting and therefore increase their generalisation properties. It may be beneficial to transfer motion information acquired using DL approaches of different application domains to pose estimation of infants (e.g., Sim2Real; Doersch & Zisserman, 2019); or, to adopt interactive machine learning techniques using feedback from the users to enable modifiable and self-improving models.

As each of the recording- and data acquisition setups and their belonging classification techniques have inherent strengths and limits (Table 3), a “method-of-choice” for automated GMA does not seem to exist. One might think of an ideal solution that combines multiple setups to complement each other. However, bearing in mind classic GMA’s non-intrusive principle and its merits of being easy-to-use, time- and cost-efficient, to scale up GMA, we must avoid sophisticated, time-consuming, or intrusive setups (e.g., combining wearable

sensors or markers with a complex video recording system requiring to configure and calibrate multiple 3D cameras). Such setups are constrained by intricate technical requirements both for data acquisition and processing. For one thing, these setups may influence the infant's motor pattern, as discussed above. For another, they may prove unsuitable for everyday clinical implementation, being especially inapt in low-resource settings. This way, we would lose the basis for realising the ultimate goal of worldwide routine application of GMA. Nonetheless, depending on the purpose of the respective automated tool, e.g., precise clinical judgement versus initial rapid screening for further referrals or diagnostics, one needs to weigh in on the recording and data acquisition setups and choose and design an appropriate combination.

Regarding current tracking techniques, state-of-the-art methods such as DeepLabCut (Mathis et al., 2018) and OpenPose (Cao, Martinez, Simon, Wei, & Sheikh, 2019) show promising results when tracking both animals and human adults. A new commercial framework, WrenchAI (WrenchAI, 2020) is reported to offer much faster and more accurate adult movement tracking than OpenPose (Gupta, 2020). Whether this holds true for young infants is an open question. As pose estimation includes skeleton constraints as additional *prior* information, it needs to be examined whether such constraints truly improve the movement detection, or whether they might not be permissive for GMA, hence hindering automated detection (Rahmati, Aamo, Stavadahl, Dragon, & Adde, 2014).

Some additional technical considerations may also improve the classification models. For example, having obtained sufficient annotated data, common practice in the field is to split the dataset into three parts, i.e., training set to update model parameters, validation set to evaluate model overfitting, and testing set to assess the classification accuracy and how well the model generalises to new data. If only a small dataset is available, data splitting will become challenging and additional strategies will be required (e.g., Beleites, Neugebauer, Bocklitz, Krafft, & Popp, 2013; Riley et al., 2020; Shahinfar, Meek, & Falzon, 2020). Furthermore, given that a considerable number of features have been extracted and presented by the different studies, whether or not to include a feature pre-selection step is still an open question, depending also on the movement detection (e.g., movement shape vs body pose estimation) and learning (e.g., supervised vs unsupervised) approaches used. Without pre-selection, a significantly higher number of variables must be explored by the classification algorithm. Finally, the most popular algorithms for movement classification are currently SVMs, Random Forests and CNNs, due to their simplicity and straightforward application for a large variety of problems. Novel algorithms have been introduced to the field of automated GMA, such as the Naive Gaussian Bayesian Surprise (NGBS), applied to calculate how much each infant's movements in a dataset deviate from a group of typically developing infants as the indicator of risk for atypical GMs (Chambers et al., 2019). Similar as in choosing the suitable sensing setups, the selection of the most appropriate algorithm is also contingent on, among others, the data acquisition approaches, the dataset characteristics, and the goal of classification and detection.

Regardless of the technological refinements, currently, automated solutions are developed to complement, but not to replace human assessment in clinical practices. Extending the machine learning technology of tracking and classifying the GMs, future computer-based

approaches with multimodal setups (e.g., motion-sensor, pressure-sensitive mat, eye-tracker) may be developed to improve human performance by actively supporting the GMA assessors in real-time across multiple training and clinical settings.

## 5. Conclusion

Automated video-based approaches, being authentic to the non-intrusive principle of the classic GMA, supported by rapid advancements in AI technologies, have the potential to scale up the clinical application of GMA. Technology advancements will enable better data pre-processing (e.g., image enhancement, noise attenuation, region-of-interest detection), improve feature extraction and analyses and lead to an objective and more accurate prediction. Currently, automated GMA models are yet inferior to human experts. Despite their classification performance, current models can deal with but a fraction of the tasks (e.g., some binary or multiple classifications) that a human expert can solve in a standard GMA of a few minutes (e.g., evaluating simultaneously the movement characteristics including complexity and variability, age-specific repertoire, posture, and motor optimality). It is, thus, unlikely that human assessors can be replaced by fully automated systems in the near future. To improve computer-based approaches, there is still a lot to learn from the human GMA experts. This concerns prerequisites for performing GMA and evaluation process embracing manifold aspects to encapsulate infant movements. While developing automated detection and classification models for GMA, a parallel line of research is needed aiming at interactive, real-time support and training for human GMA assessors. By supplementing human faculties (versatile and adaptable to complex and ever-changing situations, proficient in transferring rich experience to novel situations) with computerised tools (objective, stable, fast, and extendable), a future augmented GMA may yield outstanding performance, superior to what humans or computers could achieve alone.

While recent studies focused primarily on the prediction of CP, it is crucial for future research to look beyond this narrow field and open up to the potential of applying GMA to identify deviant early motor functions in infants with various developmental and neurological disorders, infectious diseases affecting the developing nervous system, and genetic disorders. Availing of the advanced computer-vision technology, GMA may be employed to detect more general disintegrity of the developing nervous system through fine-grained high-standard analyses of infant early motor functions. Based on the profound understanding of GMs, incorporating state-of-the-art technology, we are envisioning a worldwide daily clinical application of GMA for the youngest population in the near future.

## Acknowledgements

We would like to thank Dr. Simon Reich for his initial comments on technical issues of this manuscript. This work was supported by Bill-Melinda Gates Foundation (OPP1128871); FWF [grant number KLI 811]; the OeNB Jubiläumsfonds, the Leibniz Foundation (LSC Audacity Award); and the Volkswagen Foundation (IDENTIFIED, Deep Movement Diagnostics).

## References

Alcantarilla P, Bartoli A, & Davison A (2012, October 7). KAZE features. 10.1007/978-3-642-33783-3\_16.

- Ali J, Charman T, Johnson MH, Jones EJH, & BASIS/STAARS Team. (2020). Early motor differences in infants at elevated likelihood of autism spectrum disorder and/or attention deficit hyperactivity disorder. *Journal of Autism and Developmental Disorders*, 50(12), 4367–4384. 10.1007/s10803-020-04489-1. [PubMed: 32328858]
- Beck F, Koch S, & Weiskopf D (2016). Visual analysis and dissemination of scientific literature collections with SurVis. *IEEE Transactions on Visualization and Computer Graphics*, 22(1), 180–189. 10.1109/TVCG.2015.2467757. [PubMed: 26529698]
- Beleites C, Neugebauer U, Bocklitz T, Krafft C, & Popp J (2013). Sample size planning for classification models. *Analytica Chimica Acta*, 760, 25–33. 10.1016/j.aca.2012.11.007. [PubMed: 23265730]
- Bölte S, Bartl-Pokorny KD, Jonsson U, Berggren S, Zhang D, Kostrzewa E, ... Marschik PB (2016). How can clinicians detect and treat autism early? Methodological trends of technology use in research. *Acta Paediatrica (Oslo, Norway : 1992) Supplement*, 105(2), 137–144. 10.1111/apa.13243.
- Bosanquet M, Copeland L, Ware R, & Boyd R (2013). A systematic review of tests to predict cerebral palsy in young children. *Developmental Medicine and Child Neurology*, 55(5), 418–426. 10.1111/dmcn.12140. [PubMed: 23574478]
- Brasil P, Pereira JP, Moreira ME, Ribeiro Nogueira RM, Damasceno L, Wakimoto M, Rabello RS, Valderramos SG, Halai U-A, Salles TS, Zin AA, Horovitz D, Daltro P, Boechat M, Raja Gabaglia C, ... Nielsen-Saines K (2016). Zika virus infection in pregnant women in Rio De Janeiro. *The New England Journal of Medicine*, 375(24), 2321–2334. 10.1056/NEJMoa1602412. [PubMed: 26943629]
- Brox JM (2011). Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3), 500–513. 10.1109/TPAMI.2010.143. [PubMed: 20714020]
- Brox T, Bruhn A, Papenber N, & Weickert J (2004). High accuracy optical flow estimation based on a theory for warping. *European Conference on Computer Vision*, 25–36. 10.1007/978-3-540-24673-2\_3.
- Burger M, & Louw QA (2009). The predictive validity of general movements – A systematic review. *European Journal of Paediatric Neurology*, 13(5), 408–420. 10.1016/j.ejpn.2008.09.004. [PubMed: 19036618]
- Calinon S (2016). A tutorial on task-parameterized movement learning and retrieval. *Intelligent Service Robotics*, 9(1), 1–29. 10.1007/s11370-015-0187-9.
- Cao Z, Martinez GH, Simon T, Wei S, & Sheikh YA (2019). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 10.1109/TPAMI.2019.2929257.
- Cenci A, Liciotti D, Frontoni E, Zingaretti P, & Carnielli VP (2017). Movements analysis of preterm infants by using depth sensor. *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*, 1–9. 10.1145/3109761.3109773.
- Chambers C, Seethapathi N, Saluja R, Loeb H, Pierce S, Bogen D, ... Kording KP (2019). Computer vision to automatically assess infant neuromotor risk. *Animal behavior and cognition*. 10.1101/756262.
- Doersch C, & Zisserman A (2019). Sim2real transfer learning for 3D human pose estimation: Motion to the rescue. *Advances in Neural Information Processing Systems*, 32, 12949–12961.
- Doroniewicz I, Ledwo DJ, Affanasowicz A, Kieszczy ska K, Latos D, Matyja M, ... My liwiec A (2020). Writhing movement detection in newborns on the second and third day of life using pose-based feature machine learning classification. *Sensors*, 20(21), 5986. 10.3390/s20215986.
- Einspieler C, & Marschik PB (2019). Regression in Rett syndrome: Developmental pathways to its onset. *Neuroscience and Biobehavioral Reviews*, 98, 320–332. 10.1016/j.neubiorev.2019.01.028. [PubMed: 30832924]
- Einspieler C, & Marschik PB (2020a). The developmental spectrum of prenatal Zika virus exposure. *The Lancet Child & Adolescent Health*, 4(5), 345–346. 10.1016/S2352-4642(20)30071-7. [PubMed: 32199079]

- Einspieler C, & Marschik PB (2020b). Desideratum: A developmentalist view of Zika virus infection. *The Lancet Infectious Diseases*, 0(0). 10.1016/S1473-3099(20)30454-0.
- Einspieler C, & Prechtel HFR (2005). Prechtel'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. 10.1002/mrdd.20051. [PubMed: 15856440]
- Einspieler C, Sigafoos J, Bölte S, Bratl-Pokorny KD, Landa R, & Marschik PB (2014). Highlighting the first 5 months of life: General movements in infants later diagnosed with autism spectrum disorder or Rett syndrome. *Research in Autism Spectrum Disorders*, 8(3), 286–291. 10.1016/j.rasd.2013.12.013. [PubMed: 29770159]
- Einspieler C, Bos AF, Krieger-Tomantschger M, Alvarado E, Barbosa VM, Bertocelli N, Burger M, Chorna O, Del Secco S, DeRegnier R-A, Hüning B, Ko J, Lucaccioni L, Maeda T, Marchi V, Martín E, Morgan C, Mutlu A, Nogolová A, ... Marschik PB (2019). Cerebral palsy: Early markers of clinical phenotype and functional outcome. *Journal of Clinical Medicine*, 8(10). 10.3390/jcm8101616.
- Einspieler C, Bos AF, Libertus ME, & Marschik PB (2016). The general movement assessment helps us to identify preterm infants at risk for cognitive dysfunction. *Frontiers in Psychology*, 7, 406. 10.3389/fpsyg.2016.00406. [PubMed: 27047429]
- Einspieler C, Marschik PB, & Prechtel HFR (2008). Human motor behaviour prenatal origin and early postnatal development. *Zeitschrift für Psychologie / Journal of Psychology*, 216, 147–153. 10.1027/0044-3409.216.3.147.
- Einspieler C, Peharz R, & Marschik PB (2016). Fidgety movements – Tiny in appearance, but huge in impact. *Jornal de Pediatria*, 92, 64–70. 10.1016/j.jped.2015.12.003.
- Einspieler C, Utsch F, Brasil P, Panvequio Aizawa CY, Peyton C, Hydee Hasue R, Françoso Genovesi F, Damasceno L, Moreira ME, Adachi K, Marschik PB, Nielsen-Saines K, & GM Zika Working Group. (2019). Association of infants exposed to prenatal Zika virus infection with their clinical, neurologic, and developmental status evaluated via the general movement assessment tool. *JAMA Network Open*, 2(1), e187235. 10.1001/jamanetworkopen.2018.7235. [PubMed: 30657537]
- Groos D, Adde L, Støen R, Ramampiaro H, & Ihlen EAF (2020). Towards human performance on automatic motion tracking of infant spontaneous movements. *ArXiv:2010.05949 [Cs]*. <http://arxiv.org/abs/2010.05949>.
- Grunewaldt KH, Fjørtoft T, Bjuland KJ, Brubakk A-M, Eikenes L, Håberg AK, ... Skranes J (2014). Follow-up at age 10 years in ELBW children—Functional outcome, brain morphology and results from motor assessments in infancy. *Early Human Development*, 90(10), 571–578. 10.1016/j.earlhumdev.2014.07.005. [PubMed: 25103790]
- Gupta V (2020). Pose detection comparison: Wmchai vs OpenPose (learn OpenCV). <https://www.learnopencv.com/pose-detection-comparison-wmchai-vs-openpose/>.
- Herrero D, Einspieler C, Panvequio Aizawa CY, Mutlu A, Yang H, Nogolová A, ... GenGM Study Group. (2017). The motor repertoire in 3- to 5-month old infants with Down syndrome. *Research in Developmental Disabilities*, 67, 1–8. 10.1016/j.ridd.2017.05.006. [PubMed: 28586709]
- Hesse N, Pujades S, Romero J, Black MJ, Bodensteiner C, Arens M, ... Sebastian Schroeder A (2018). Learning an infant body model from RGB-D data for accurate full body motion analysis. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 (Vol. 11070, Pp. 792–800)*. 10.1007/978-3-030-00928-1\_89.
- Hesse N, Schroeder S, Müller-Felber W, Bodensteiner C, Arens M, & Hofmann G (2018). Markerless motion analysis for early detection of infantile movement disorders. In Eskola H, Väisänen O, Viik J, & Hyttinen J (Eds.), *EMBECC & NBC 2017 (vol. 65, pp. 197–200)*. Singapore: Springer. 10.1007/978-981-10-5122-7\_50.
- Hyde KK, Novack MN, LaHaye N, Parlett-Pelleriti C, Anden R, Dixon DR, ... Linstead E (2019). Applications of supervised machine learning in autism spectrum disorder research: A review. *Review Journal of Autism and Developmental Disorders*, 6(2), 128–146. 10.1007/s40489-019-00158-x.
- Ihlen EAF, Støen R, Boswell L, de Regnier R-A, Fjørtoft T, Gaebler-Spira D, ... Adde L (2020). 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. 10.3390/jcm9010005.

- Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P, & Schaal S (2013). Dynamical movement primitives: Learning attractor models for motor behaviors. *Neural Computation*, 25(2), 328–373. 10.1162/NECO\_a\_00393. [PubMed: 23148415]
- Irshad MT, Nisar MA, Gouverneur P, Rapp M, & Grzegorzec M (2020). AI approaches towards Prechtl's assessment of general movements: A systematic literature review. *Sensors*, 20(18), 5321. 10.3390/s20185321.
- Khansari-Zadeh SM, & Billard A (2011). Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5), 943–957. 10.1109/TRO.2011.2159412.
- Kuka ka J, Golkov V, & Cremers D (2017). Regularization for deep learning: A taxonomy. ArXiv:1710.10686 [Cs, Stat]. <http://arxiv.org/abs/1710.10686>.
- Kwong AKL, Fitzgerald TL, Doyle LW, Cheong JLY, & Spittle AJ (2018). Predictive validity of spontaneous early infant movement for later cerebral palsy: A systematic review. *Developmental Medicine and Child Neurology*, 60(5), 480–489. 10.1111/dmcn.13697. [PubMed: 29468662]
- Lin G, Zhu H, Fan Y, & Fan C (2011). Human action recognition based on random spectral regression. In Deng H, Miao D, Lei J, & Wang FL (Eds.), *Artificial intelligence and computational intelligence* (pp. 451–461). Springer. 10.1007/978-3-642-23896-3\_56.
- Marcroft C, Khan A, Embleton ND, Trenell M, & Plötz T (2015). Movement recognition technology as a method of assessing spontaneous general movements in high risk infants. *Frontiers in Neurology*, 5. 10.3389/fneur.2014.00284.
- Marschik PB, Einspieler C, Sigafoos J, Enzinger C, & Bölte S (2016). The interdisciplinary quest for behavioral biomarkers pinpointing developmental disorders. *Developmental Neurorehabilitation*, 19(2), 73–74. 10.3109/17518423.2014.916362. [PubMed: 24840633]
- Marschik PB, Pokorny FB, Peharz R, Zhang D, O'Muircheartaigh J, Roeyers H, Bölte S, Spittle AJ, Urlesberger B, Schuller B, Poustka L, Ozonoff S, Pernkopf F, Pock T, Tammimies K, Enzinger C, Krieger M, Tomantschger I, Bartl-Pokorny KD, ... Kaufmann WE (2017). A novel way to measure and predict development: A heuristic approach to facilitate the early detection of neurodevelopmental disorders. *Current Neurology and Neuroscience Reports*, 17(5). 10.1007/s11910-017-0748-8.
- Marschik PB, Soloveichick M, Windpassinger C, & Einspieler C (2015). General movements in genetic disorders: A first look into Cornelia de Lange syndrome. *Developmental Neurorehabilitation*, 18(4), 280–282. 10.3109/17518423.2013.859180. [PubMed: 24304241]
- Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, ... Bethge M (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), 1281–1289. 10.1038/s41593-018-0209-y. [PubMed: 30127430]
- McDonald NM, Senturk D, Scheffler A, Brian JA, Carver LJ, Charman T, ... Jeste SS (2020). Developmental trajectories of infants with multiplex family risk for autism: A baby siblings research consortium study. *JAMA Neurology*, 77(1), 73–81. 10.1001/jamaneurol.2019.3341. [PubMed: 31589284]
- Moccia S, Migliorelli L, Carnielli V, & Frontoni E (2020). Preterm infants' pose estimation with spatio-temporal features. *IEEE Transactions on Biomedical Engineering*. 10.1109/TBME.2019.2961448, 1–1.
- Murphy D, & Spooen W (2012). EU-AIMS: A boost to autism research. *Nature Reviews Drug Discovery*, 11, 815–816. 10.1038/nrd3881.
- Novak I, Morgan C, Adde L, Blackman J, Boyd RN, Brunstrom-Hernandez J, Cioni G, Damiano D, Darrah J, Eliasson A-C, de Vries LS, Einspieler C, Fahey M, Fehlings D, Ferriero DM, Fetters L, Fiori S, Forssberg H, Gordon AM, ... Badawi N (2017). Early, accurate diagnosis and early intervention in cerebral palsy: Advances in diagnosis and treatment. *JAMA Pediatrics*, 171(9), 897–907. 10.1001/jamapediatrics.2017.1689. [PubMed: 28715518]
- Orlandi S, Raghuram K, Smith CR, Mansueto D, Church P, Shah V, ... Chau T (2018). Detection of atypical and typical infant movements using computer-based video analysis. *IEEE Engineering in Medicine and Biology Society. Annual Conference, 2018*, 3598–3601. 10.1109/EMBC.2018.8513078.
- Ouss L, Le Normand M-T, Bailly K, Leitgel Gille M, Gosme C, Simas R, ... Guergova-Kuras M (2018). Developmental trajectories of hand movements in typical infants and those at risk of

- developmental disorders: An observational study of kinematics during the first year of life. *Frontiers in Psychology*, 9. 10.3389/fpsyg.2018.00083.
- Ouss L, Palestra G, Saint-Georges C, Leitgel Gille M, Afshar M, Pellerin H, ... Cohen D (2020). Behavior and interaction imaging at 9 months of age predict autism/intellectual disability in high-risk infants with West syndrome. *Translational Psychiatry*, 10(1), 1–7. 10.1038/s41398-020-0743-8. [PubMed: 32066695]
- Ozonoff S, Young GS, Landa RJ, Brian J, Bryson S, Charman T, ... Iosif A-M (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 56(9), 988–998. 10.1111/jcpp.12421.
- Pal NR, & Pal SK (1993). A review on image segmentation techniques. *Pattern Recognition*, 26(9), 1277–1294. 10.1016/0031-3203(93)90135-J.
- Palchik AB, Einspieler C, Evstafeyeva IV, Talisa VB, & Marschik PB (2013). Intra-uterine exposure to maternal opiate abuse and HIV: The impact on the developing nervous system. *Early Human Development*, 89(4), 229–235. 10.1016/j.earlhumdev.2013.02.004. [PubMed: 23490656]
- Paraschos A, Daniel C, Peters J, & Neumann G (2018). Using probabilistic movement primitives in robotics. *Autonomous Robots*, 42(3), 529–551. 10.1007/s10514-017-9648-7.
- Peyton C, & Einspieler C (2018). General movements: A behavioral biomarker of later motor and cognitive dysfunction in NICU graduates. *Pediatric Annals*, 47(4), e159–e164. 10.3928/19382359-20180325-01. [PubMed: 29668025]
- Prechtl HFR (1990). Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Human Development*, 23(3), 151–158. 10.1016/0378-3782(90)90011-7. [PubMed: 2253578]
- Prechtl HFR, Einspieler C, Cioni G, Bos AF, Ferrari F, & Sontheimer D (1997). An early marker for neurological deficits after perinatal brain lesions. *Lancet (London, England)*, 349(9062), 1361–1363. 10.1016/S0140-6736(96)10182-3.
- Rahmati H, Aamo OM, Stavdahl Ø, Dragon R, & Adde L (2014). Video-based early cerebral palsy prediction using motion segmentation. *IEEE Engineering in Medicine and Biology Society. Annual Conference, 2014*, 3779–3783. 10.1109/EMBC.2014.6944446.
- Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, ... Smeden M (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, 368. 10.1136/bmj.m441.
- Romeo DMM, Guzzetta A, Scoto M, Cioni M, Patusi P, Mazzone D, ... Romeo MG (2008). Early neurologic assessment in preterm-infants: Integration of traditional neurologic examination and observation of general movements. *European Journal of Paediatric Neurology*, 12(3), 183–189. 10.1016/j.ejpn.2007.07.008. [PubMed: 17881261]
- Salavati S, Einspieler C, Vagelli G, Zhang D, Pansy J, Burgerhof JGM, ... Bos AF (2017). The association between the early motor repertoire and language development in term children born after normal pregnancy. *Early Human Development*, 111, 30–35. 10.1016/j.earlhumdev.2017.05.006. [PubMed: 28549271]
- Schmidt W, Regan M, Fahey M, & Paplinski A (2019). General movement assessment by machine learning: Why is it so difficult? *Journal of Medical Artificial Intelligence*, 2. 10.21037/jmai.2019.06.02, 2–2.
- Shahinfar S, Meek P, & Falzon G (2020). “How many images do I need?” understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecological Informatics*, 57, Article 101085. 10.1016/j.ecoinf.2020.101085.
- Shephard E, Bedford R, Milosavljevic B, Gliga T, Jones EJH, Pickles A, ... BASIS Team. (2019). Early developmental pathways to childhood symptoms of attention-deficit hyperactivity disorder, anxiety and autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 60(9), 963–974. 10.1111/jcpp.12947.
- Silva N, Blaschek T, Jianu R, Rodrigues N, Weiskopf D, Raubal M, ... Schreck T (2019). Eye tracking support for visual analytics systems: Theoretical foundations, opportunities, and research themes. *Proceedings of the 2019 ACM Symposium on Eye Tracking Research & Applications*, 131–139. 10.1145/3314111.3319919.

- Silva N, Schreck T, Veas E, Sabol V, Eggeling E, & Fellner DW (2018). Leveraging eye-gaze and time-series features to predict user interests and build a recommendation model for visual analysis. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. 10.1145/3204493.3204546.
- Soares-Marangoni D, de A, Tedesco NM, Nascimento AL, Almeida PRD, & Santos Pereira CND (2019). General movements and motor outcomes in two infants exposed to Zika virus: Brief report. *Developmental Neurorehabilitation*, 22(1), 71–74. 10.1080/17518423.2018.1437843. [PubMed: 29452026]
- Soleimani F, Teymouri R, & Biglarian A (2013). Predicting developmental disorder in infants using an artificial neural network. *Acta Medica Iranica*, 51(6), 347–352. [PubMed: 23852837]
- Spittle AJ, Olsen J, Kwong A, Doyle LW, Marschik PB, Einspieler C, ... Cheong J (2016). The baby moves prospective cohort study protocol: Using a smartphone application with the general movements assessment to predict neurodevelopmental outcomes at age 2 years for extremely preterm or extremely low birthweight infants. *BMJ Open*, 6(10), e013446. 10.1136/bmjopen-2016-013446.
- Tan C, Sun F, Kong T, Zhang W, Yang C, & Liu C (2018). A survey on deep transfer learning. ArXiv:1808.01974 [Cs, Stat]. <http://arxiv.org/abs/1808.01974>.
- Tomantschger I, Herrero D, Einspieler C, Hamamura C, Voos MC, & Marschik PB (2018). The general movement assessment in non-European low- and middle-income countries. *Revista de Saude Publica*, 52, 6. 10.11606/S1518-8787.2018052000332. [PubMed: 29412374]
- Valle SC, Støen R, Sæther R, Jensenius AR, & Adde L (2015). Test–retest reliability of computer-based video analysis of general movements in healthy term-born infants. *Early Human Development*, 91(10), 555–558. 10.1016/j.earlhumdev.2015.07.001. [PubMed: 26217934]
- Varcin KJ, & Nelson CA (2016). A developmental neuroscience approach to the search for biomarkers in autism spectrum disorder. *Current Opinion in Neurology*, 29(2), 123–129. 10.1097/WCO.0000000000000298. [PubMed: 26953849]
- WrnchAI. (Accessed: 2020-08-17). <https://wrnch.ai/>.
- Yuge M, Marschik PB, Nakajima Y, Yamori Y, Kanda T, Hirota H, ... Einspieler C (2011). Movements and postures of infants aged 3 to 5 months: To what extent is their optimality related to perinatal events and to the neurological outcome? *Early Human Development*, 87(3), 231–237. 10.1016/j.earlhumdev.2010.12.046. [PubMed: 21295927]
- Zang F-F, Yang H, Han Q, Cao J-Y, Tomantschger I, Kriebler M, ... Einspieler C (2016). Very low birth weight infants in China: The predictive value of the motor repertoire at 3 to 5 months for the motor performance at 12 months. *Early Human Development*, 100, 27–32. 10.1016/j.earlhumdev.2016.03.010. [PubMed: 27391870]
- Zappella M, Einspieler C, Bartl-Pokorny KD, Kriebler M, Coleman M, Bölte S, ... Marschik PB (2015). What do home videos tell us about early motor and socio-communicative behaviours in children with autistic features during the second year of life—An exploratory study. *Early Human Development*, 91(10), 569–575. 10.1016/j.earlhumdev.2015.07.006. [PubMed: 26246137]
- Zotero (about) (Accessed: 2020-08-07). <https://www.zotero.org/about/>.

## Further reading

- Adde L, Helbostad JL, Jensenius AR, Langaas M, & Støen R (2013). Identification of fidgety movements and prediction of CP by the use of computer-based video analysis is more accurate when based on two video recordings. *Physiotherapy Theory and Practice*, 29(6), 469–475. 10.3109/09593985.2012.757404. [PubMed: 23343036]
- Adde L, Helbostad JL, Jensenius AR, Taraldsen G, Grunewaldt KH, & Støen R (2010). Early prediction of cerebral palsy by computer-based video analysis of general movements: A feasibility study: Early Computer-based Prediction of CP. *Developmental Medicine & Child Neurology*, 52(8), 773–778. 10.1111/j.1469-8749.2010.03629.x. [PubMed: 20187882]
- Adde L, Yang H, Sæther R, Jensenius AR, Ihlen E, Cao J, & Støen R (2018). Characteristics of general movements in preterm infants assessed by computer-based video analysis. *Physiotherapy Theory and Practice*, 34(4), 286–292. 10.1080/09593985.2017.1391908. [PubMed: 29064734]

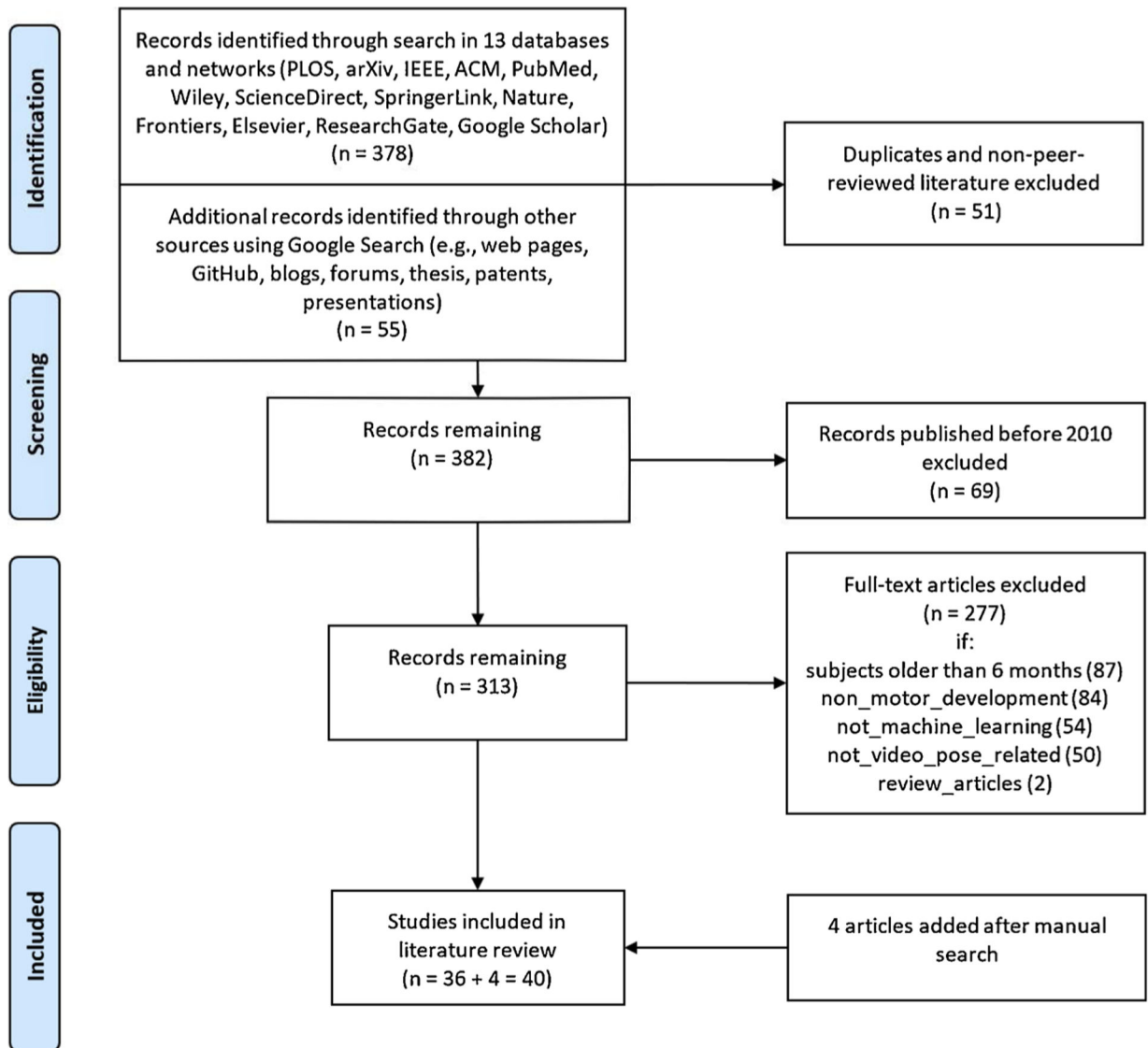


- Baccinelli W, Bulgheroni M, Simonetti V, Fulceri F, Caruso A, Gila L, & Scattoni ML (2020). Movidea: A software package for automatic video analysis of movements in infants at risk for neurodevelopmental disorders. *Brain Sciences*, 10(4). 10.3390/brainsci10040203.
- Caruso A, Gila L, Fulceri F, Salvitti T, Micai M, Baccinelli W, ... Scattoni ML (2020). Early motor development predicts clinical outcomes of siblings at high-risk for autism: Insight from an innovative motion-tracking technology. *Brain Sciences*, 10(6), 379. 10.3390/brainsci10060379.
- Dai X, Wang S, Li H, Yue H, & Min J (2019). Image-assisted discrimination method for neurodevelopmental disorders in infants based on multi-feature fusion and ensemble learning. In Liang P, Goel V, & Shan C (Eds.), *Brain Informatics* (pp. 105–114). Springer International Publishing. 10.1007/978-3-030-37078-7\_11.
- Das D, Fry K, & Howard AM (2018). Vision-based detection of simultaneous kicking for identifying movement characteristics of infants at-risk for neuro-disorders. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1413–1418). 10.1109/ICMLA.2018.00230.
- Einspieler C, Prechtel HRF, Bos A, Ferrari F, & Cioni G (2004). *Prechtel's method on the qualitative assessment of general movements in preterm, term and young infants* (1st edition). Mac Keith Press.
- Gajniyarov I, Mikhailov I, Starodubtsev I, Obabkov I, Lvova O, Suleymanova E, & Antipina I (2019). The motion capture as behavior analyzing method of spontaneous motor activity in human infants. 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON) (pp. 0681–0684). 10.1109/SIBIRCON48586.2019.8958041.
- Hesse N, Schroder AS, Muller-Felber W, Bodensteiner C, Arens M, & Hofmann UG (2017). Body pose estimation in depth images for infant motion analysis. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1909–1912). 10.1109/EMBC.2017.8037221.
- Hesse N, Stachowiak G, Breuer T, & Arens M (2015). Estimating body pose of infants in depth images using random ferns. 2015 IEEE International Conference on Computer Vision Workshop (ICCVW) (pp. 427–435). 10.1109/ICCVW.2015.63.
- Hesse N, Bodensteiner C, Arens M, Hofmann UG, Weinberger R, & Sebastian Schroeder A (2019). Computer vision for medical infant motion analysis: State of the art and RGB-D data set. In Leal-Taixé L & Roth S (Eds.), *Computer Vision – ECCV 2018 Workshops* (pp. 32–49). Springer International Publishing. 10.1007/978-3-030-11024-6\_3.
- Hesse N, Pujades S, Black M, Arens M, Hofmann U, & Schroeder S (2019). Learning and tracking the 3D body shape of freely moving infants from RGB-D sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. 10.1109/TPAMI.2019.2917908.
- Machireddy A, van Santen J, Wilson JL, Myers J, Hadders-Algra M, & Song X (2017). A video/IMU hybrid system for movement estimation in infants. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, 2017*, (pp. 730–733). 10.1109/EMBC.2017.8036928.
- Marchi V, Belmonti V, Cecchi F, Coluccini M, Ghirri P, Grassi A, ... Guzzetta A (2020). Movement analysis in early infancy: Towards a motion biomarker of age. *Early Human Development*, 142, 104942. 10.1016/j.earlhumdev.2019.104942. [PubMed: 32028093]
- Marchi V, Hakala A, Knight A, D'Acunto F, Scattoni ML, Guzzetta A, & Vanhatalo S (2019). Automated pose estimation captures key aspects of general movements at eight to 17 weeks from conventional videos. *Acta Paediatrica*, 108(10), 1817–1824. 10.1111/apa.14781. [PubMed: 30883894]
- McCay K, Ho E, Marcroft C, & Embleton ND (2019). Establishing pose based features using histograms for the detection of abnormal infant movements. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 5469–5472). 10.1109/EMBC.2019.8857680.
- McCay K, Ho E, Shum HPH, Fehringer G, Marcroft C, & Embleton ND (2020). Abnormal infant movements classification with deep learning on pose-based features. *IEEE Access*, 8, 51582–51592. 10.1109/ACCESS.2020.2980269.

- Moccia S, Migliorelli L, Pietrini R, & Frontoni E (2019). Preterm infants' limb-pose estimation from depth images using convolutional neural networks. 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 1–7. 10.1109/CIBCB.2019.8791242.
- Olsen MD (2016). Motion tracking of infants in risk of cerebral palsy [PhD Thesis]. Technical University of Denmark. DTU Compute PHD-2015.
- Raghuram K, Orlandi S, Shah V, Chau T, Luther M, Banihani R, & Church P (2019). Automated movement analysis to predict motor impairment in preterm infants: A retrospective study. *Journal of Perinatology*, 39(10), 1362–1369. 10.1038/s41372-019-0464-0. [PubMed: 31431653]
- Rahmati H, Dragon R, Aamo OM, Adde L, Stavdahl Ø, & Van Gool L (2015). Weakly supervised motion segmentation with particle matching. *Computer Vision and Image Understanding*, 140, 30–42. 10.1016/j.cviu.2015.07.004.
- Rahmati H, Martens H, Aamo OM, Stavdahl O, Stoen R, & Adde L (2016). Frequency analysis and feature reduction method for prediction of cerebral palsy in young infants. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 24(11), 1225–1234. 10.1109/TNSRE.2016.2539390.
- Rahmati H, Dragon R, Aamo OM, van Gool L, & Adde L (2014). Motion segmentation with weak labeling priors. In Jiang X, Hornegger J, & Koch R (Eds.), *Pattern Recognition* (pp. 159–171). Springer International Publishing. 10.1007/978-3-319-11752-2\_13.
- Schroeder AS, Hesse N, Weinberger R, Tacke U, Gerstl L, Hilgendorff A, ... Hadders-Algra M (2020). General movement assessment from videos of computed 3D infant body models is equally effective compared to conventional RGB video rating. *Early Human Development*, 144, 104967. 10.1016/j.earlhumdev.2020.104967. [PubMed: 32304982]
- Serrano MM, Yu-Ping C, Howard A, & Vela PA (2016). Lower limb pose estimation for monitoring the kicking patterns of infants. *IEEE Engineering in Medicine and Biology Society. Annual Conference, 2016*, 2157–2160. 10.1109/EMBC.2016.7591156.
- Shivakumar SS, Loeb H, Bogen DK, Shofer F, Bryant P, Prosser L, & Johnson MJ (2017). Stereo 3D tracking of infants in natural play conditions. *International Conference on Rehabilitation Robotics, 2017*, 841–846. 10.1109/ICORR.2017.8009353.
- Stahl A, Schellewald C, Stavdahl Ø, Aamo OM, Adde L, & Kirkerød H (2012). An optical flow-based method to predict infantile cerebral palsy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 20(4), 605–614. 10.1109/TNSRE.2012.2195030.
- Støen R, Songstad NT, Silberg IE, Fjørtoft T, Jensenius AR, & Adde L (2017). Computer-based video analysis identifies infants with absence of fidgety movements. *Pediatric Research*, 82(4), 665–670. 10.1038/pr.2017.121. [PubMed: 28745715]
- Tsuji T, Nakashima S, Hayashi H, Soh Z, Furui A, Shibasaki T, ... Shimatani K (2020). Markerless measurement and evaluation of general movements in infants. *Scientific Reports*, 10(1), 1422. 10.1038/s41598-020-57580-z. [PubMed: 31996716]

**What this paper adds?**

- An overview of computer vision-based approaches in the study of general movements is provided.
- The advantages, limitations, and future directions of vision-based approaches in performing automated general movement assessment (GMA) are discussed.
- Prospects of computer-driven GMA are discussed. The necessity of understanding the nature of general movements and GMA while developing automated solutions is highlighted.
- It is suggested that future research shall look beyond the narrow field of detecting cerebral palsy and open up to the potential of applying GMA to identify more general disintegrity of the developing nervous system in early infancy.



**Fig. 1.** Literature search and screening procedure.

**Table 1**

Search categories and terms.

Categories	Search Terms
<b>General Movements</b>	baby OR child* OR infan* OR newborn OR *term OR neonatal OR abnormal OR anomaly OR atypical OR disorder OR risk OR sign* OR typical OR diagnos* OR analys* OR early OR assessment OR behavio* OR *marker OR cerebral palsy OR development* OR fidgety OR body OR gma OR gm OR outcome OR general movement* OR motor AND
<b>Machine Learning</b>	accuracy OR adaptive OR advanced OR auto OR biosensor OR classification OR detect* OR learn* predict* OR recognition OR recommend* OR sens* OR neuro* OR algorithm* OR deep OR model* OR machine AND
<b>Computer Vision</b>	2d OR 3d OR action OR activity OR classifier OR estimat* OR framework OR human OR intelligen* OR motion OR pose OR predict OR tracking OR video

Table 2

Video-based technological approaches for studying GMs.

ID	REFERENCE*	CAMERA TYPE; TECHNIQUE; PREDICTIONS	FEATURES (F): QUANTITY / DESCRIPTION; BODY PARTS (BP)	CLASSIFICATION METHODS	AGE; SAMPLE SIZE	ACCURACY (A); PRECISION (P); RECALL (R); SPECIFICITY (S)	
<i>CONVENTIONAL MACHINE LEARNING CLASSIFICATION</i>							
1	Baccinelli et al., 2020	<i>2D Video; Movement Detection; Extract movement features</i>	<i>F: 7/trajectory, motion, image; BP: hands and feet</i>	<i>Extraction of quantitative measures</i>	<i>39–41 weeks (GA); 300 videos (90 infants high risk ASD)</i>	<i>NR (ICC: 87–98)</i>	
2	Caruso et al., 2020				<i>103 videos (53 low risk, 50 high risk ASD)</i>		
3	Doroniewicz et al., 2020	<i>2D Video; Pose Estimation; Classify WMs and PR</i>	<i>F: 16/scope, nature, and location of each limb's movement; BP: limbs</i>	<i>SVM-RBF, RF, LDA</i>	<i>38–42 weeks (GA); 31 videos</i>	<i>(SVM): A:80;P:64;R:71;S:83 (RF): A:81;P:53;R:44;S:93 (DA): A:80;P:50;R:40;S:94</i>	
4**	Tsuji et al., 2020	<i>2D Video; Movement Detection; Classify GMs</i>	<i>F: 25/movement magnitude, balance, rhythm, body centre; BP: limbs</i>	<i>LLGMN</i>	<i>25–40 weeks (GA), 0–15 weeks (PTA), NR for half of the infants; 47 videos (21 infants)</i>	<i>A:91;P:NR;R:NR;S:NR</i>	
5	Schroeder et al., 2020	<i>RGB-D Video; Shape and Pose Estimation; Classify GMs</i>	<i>F: 6890/SMIL; BP: 23 joints</i>	<i>RGB-D, 3D SMIL (Auto-Generated)</i>	<i>2–4 months (PTA); 29 videos (high risk CP)</i>	<i>A:80;P:NR;R:NR;S:92</i>	
6	Hesse et al., 2019				<i>Custom Model</i>		<i>2–4 months (PTA); 12 videos</i>
7	Hesse, Boden- steiner, et al., 2019						<i>2–4 months (PTA); 136 videos (37 infants)</i>
8	Hesse, et al., 2018						
9	Hesse, Schroeder, et al., 2018						
10	Hesse et al., 2017						
11	Hesse et al., 2015						
12	Ihlen et al., 2020	<i>2D Video; Movement Detection CIMA (MEMD); Predict CP</i>	<i>F: 990/Optical Flow, BP: head, trunk, limbs</i>	<i>LDA</i>	<i>9–15 weeks (PTA); 377 videos (high-risk CP)</i>	<i>A:93;P:NR;R:NR;S:82</i>	
13	Adde et al., 2018	<i>2D Video; Movement Detection;</i>	<i>F: NR/spatial (no temporal), C<sub>SD</sub>;</i>	<i>LR, Variability of C<sub>SD</sub></i>	<i>3–5, 10–15 weeks (PTA);</i>	<i>NR (C<sub>SD</sub> is 7.5% lower during FMs in comparison to the WMs period)</i>	

ID	REFERENCE*	CAMERA TYPE; TECHNIQUE; PREDICTIONS	FEATURES (F): QUANTITY / DESCRIPTION; BODY PARTS (BP)	CLASSIFICATION METHODS	AGE; SAMPLE SIZE	ACCURACY (A); PRECISION (P); RECALL (R); SPECIFICITY (S)
		<i>Quantify FMs vs WMs, Classify GMs</i>	<i>BP: head, trunk, limbs</i>		<i>54 videos (27 infants preterm)</i>	
14	Støen et al., 2017	<i>2D Video; Movement Detection; Detect FMs</i>	<i>F: NR/spatial and temporal, C<sub>SD</sub>; BP: neck, trunk, limbs</i>	<i>Variability of C<sub>SD</sub></i>	<i>10–15 weeks (PTA); 241 videos (150 infants: 48 abnormal)</i>	<i>NR (C<sub>SD</sub> varies between R:80; S:80–90)</i>
15	Rahmati et al., 2016	<i>2D Video; Movement Detection; Predict CP</i>	<i>F: NR/Optical Flow, FFT; BP: hands, feet, head, trunk, arms</i>	<i>SVM, MRF, Particle Matching</i>	<i>2–4 month (PTA); 78 videos (78 infants: 14 CP)</i>	<i>(SVM) A:91;P:NR;R:86;S:92</i>
16	Rahmati et al., 2015	<i>2D Video; Movement Detection; Predict CP</i>	<i>F: NR/Optical Flow; BP: hands, feet, head, trunk, arms</i>			<i>(SVM) A:87;P:NR;R:NR;S:NR</i>
17	Rahmati, Amo, et al., 2014	<i>2D Video; Movement Detection; Predict CP</i>	<i>F: NR/LDOF, graph-cut; BP: hands, feet, head, trunk</i>	<i>SVM, MRF</i>		<i>A:87;P:NR;R:50;S:95</i>
18	Rahmati, Dragon, et al., 2014	<i>2D Video; Movement Detection; Predict CP</i>	<i>F: NR/LDOF, graph-cut; BP: hands, feet, head, trunk</i>		<i>A:NR;P:96;R:NR;S:NR</i>	
19	Adde et al., 2013	<i>2D Video; Movement Detection; Detect FMs, Predict CP</i>	<i>F: NR/motion, C<sub>s</sub>, Q<sub>mean</sub>, Q<sub>sd</sub> CPP; BP: neck, trunk, limbs</i>	<i>CPP</i>	<i>9–17 weeks (PTA); 104 videos (52 infants: 24M, 28F)</i>	<i>(FMs) A:NR;P:NR;R:89;S:79 (CPP) A:NR;P:NR;R:89;S:74</i>
20	Stahl et al., 2012	<i>2D Video; CIMA; Detect FMs, Predict CP</i>	<i>F: 3/Optical Flow (GPU), wavelet, spatio-temporal; BP: head, limbs</i>	<i>SVM</i>	<i>10–15 weeks (PTA); 136 videos (82 infants: 15 atypical, 67 typical)</i>	<i>A:96;P:NR;R:88;S:98</i>
21	Adde et al., 2010	<i>2D Video; Movement Detection; Predict CP</i>	<i>F: NR/ CPP, C<sub>SD</sub>, V<sub>SD</sub>, A<sub>SD</sub>, Q<sub>mean</sub>, Q<sub>median</sub>, Q<sub>SD</sub>; BP: neck, trunk, limbs</i>	<i>CPP</i>	<i>10–15 weeks (PTA); 30 videos (high-risk: 13M, 17F)</i>	<i>(CPP) A:NR;P:NR;R:85;S:88</i>
22	Marchi et al., 2020	<i>SMART-D Video (10 cameras + markers); Movement Detection; Correlate FMs age with other measures</i>	<i>F: NR/coordination, distance, global movement quality; BP: hands and feet</i>	<i>Custom Model</i>	<i>9–20 weeks (PTA); 8 videos</i>	<i>NR (Regression, R<sup>2</sup>:97)</i>
23	Marchi et al., 2019	<i>2D Video; Pose Estimation; Classify GMs</i>	<i>F: NR/OpenPose; BP: 25 joints</i>	<i>Extraction of quantitative measures</i>	<i>8–17 weeks (PTA); 21 videos (14 typical, 7 atypical)</i>	
24**	Chambers et al., 2019	<i>2D Video; Pose Estimation; Estimate risk</i>	<i>F: 38/OpenPose and kinematics, NGBS; BP: 25 joints</i>	<i>Naive Bayes, Kinematics Data</i>	<i>4–11 months; 104 videos: 85 Youtube, 19 clinical</i>	<i>A:NR;P:92;R:94;S:NR</i>
25	Dai et al., 2019	<i>2D Video; Movement</i>	<i>F: NR/Wavelet, PCA;</i>	<i>SVM, XGBoost</i>	<i>10–12 weeks (PTA); 120 videos</i>	<i>A:93;P:NR;R:95;S:92</i>

ID	REFERENCE*	CAMERA TYPE; TECHNIQUE; PREDICTIONS	FEATURES (F): QUANTITY / DESCRIPTION; BODY PARTS (BP)	CLASSIFICATION METHODS	AGE; SAMPLE SIZE	ACCURACY (A); PRECISION (P); RECALL (R); SPECIFICITY (S)
		<i>Detection; typical vs atypical</i>	<i>BP: neck, trunk, limbs</i>		<i>(60 typical, 60 atypical)</i>	
26	Gajniyarov et al., 2019	<i>2D Video Movement Detection; Analyse GMs</i>	<i>F: NR/ segmentation, wavelet, limb speed; BP: hands and feet</i>	<i>Data Pre-processing</i>	<i>10 weeks (PTA); 18 videos</i>	<i>NR (study on data preprocessing)</i>
27	Raghuram et al., 2019	<i>2D Video; Movement Detection; Detect atypical</i>	<i>F: 289/skin model, LDOF; BP: neck, trunk, limbs</i>	<i>Logistic Regression</i>	<i>3–5 months (PTA); 152 videos</i>	<i>A:66;P:NR;R:79;S:63</i>
28	Orlandi et al., 2018		<i>F: 643/skin model, LDOF; BP: neck, trunk, limbs</i>	<i>AdaBoost, Random Forest</i>	<i>3–5 months (PTA); 127 videos (98 typical, 29 atypical)</i>	<i>A:92;P:NR;R:44;S:88</i>
29**	Das et al., 2018	<i>2D Video; Movement Detection; Detect kicks</i>	<i>F: 5/KAZE, legs in same y-direction; BP: lower limbs</i>	<i>SVM</i>	<i>4–7 months (PTA); 16 videos</i>	<i>A:91;P:88;R:85;S:NR</i>
30	Cenci et al., 2017	<i>RGB-D Video; Movement Detection; Probability of change</i>	<i>F: 10/velocity, acceleration amplitude, volume; BP: limbs</i>	<i>K-means, Markov Chains</i>	<i>37–38 weeks (GA); 35 videos (1 infant)</i>	<i>NR (initial test-phase)</i>
31	Machireddy et al., 2017	<i>2D Video; Movement Detection; Detect FMs</i>	<i>F: NR/sensor fusion, EKF; BP: limbs</i>	<i>SVM</i>	<i>2–4 months; 20 videos</i>	<i>A:84;P:NR;R:NR;S:NR</i>
32	Marschik et al., 2017	<i>2D Video; Multimodal Detection; NA</i>	<i>F: NR/multimodal fusion; BP: the whole body</i>	<i>Heuristic</i>	<i>0–4 months; NA</i>	<i>NA</i>
33**	Shivakumar et al., 2017	<i>RGB-D Video; Movement Detection; Track Body Attributes</i>	<i>F: NR/Optical Flow; BP: limbs</i>	<i>Adaptive Window, K-means</i>	<i>3–11 months (PTA); 3 videos (typical)</i>	<i>A:NR;P:NR;R:NR;S:NR</i>
34**	Serrano et al., 2016	<i>RGB-D Video; Pose Estimation; Kicking Patterns Analysis</i>	<i>F: NR/lower limb pose, RPSR; BP: lower limbs</i>	<i>Kicking Patterns of Robot</i>	<i>NR; 1 robotic infant</i>	<i>NR (qualitative analysis)</i>
35**	Olsen, 2015	<i>RGB-D Video; Pose Estimation; Detect Kickings</i>	<i>F: NR/Optical Flow; BP: stomach, head, limbs, feet</i>	<i>K-NN, Classification Tree, SVM</i>	<i>1–6 months; 11 videos</i>	<i>A:90;P:NR;R:NR;S:NR</i>
<b>DEEP LEARNING CLASSIFICATION</b>						
36	McCay et al., 2020	<i>2D Video; Pose Estimation; Classify GMs</i>	<i>F: NR/OpenPose, HOJO2D, HOJD2D; BP: 14 joints</i>	<i>FCNet model</i>	<i>2–4 months (PTA); 12 videos</i>	<i>A:NR;P:NR;R:NR;S:NR</i>
37	McCay et al., 2019					
38	Moccia et al., 2020	<i>RGB-D Video; Pose Estimation; Detect Joints</i>	<i>F: NR/spatio- temporal; BP: shoulders, elbows, wrists, hips, knees, ankles</i>	<i>Dual CNNs</i>	<i>31–36 weeks (GA); 16 videos</i>	<i>A:NR;P:NR;R:NR;S:NR</i>
39	Moccia et al., 2019					



ID	REFERENCE*	CAMERA TYPE; TECHNIQUE; PREDICTIONS	FEATURES (F); QUANTITY / DESCRIPTION; BODY PARTS (BP)	CLASSIFICATION METHODS	AGE; SAMPLE SIZE	ACCURACY (A); PRECISION (P); RECALL (R); SPECIFICITY (S)
40	Schmidt et al., 2019	2D Video; Movement Detection; Classify GMs	F: NR/OpticalFlow, FFT, Keras VGG19; BP: limbs	LSTM	2-4 month (PTA); 78 videos (78 infants; 14 CP)	A:65;P:NR;R:51;S:27

\* Articles are first arranged in descending order of the publication year, followed by ascending order of the last name of the first author. Studies with an inherent connection, i.e., leading authors are identical or worked jointly, are stacked together and shaded with the same background colour, also ordered first by the publication year and then by the last name of the first author.

\*\* Studies in which the ages of the participants fell (partly) beyond the appropriate range according to the standard GMA (Einspieler et al., 2014), or the age range was (partly) missing.

Key of Terms.

*Generic:* ASD – Autism Spectrum Disorder; CP – Cerebral Palsy; CS – Cramped Synchronised; FM – Fidgety Movements; GA – Gestational Age; GMS – General Movements; GMA – General Movement Assessment; NA – Not Applicable; NR – Not Reported; PTA – Postterm age. PR – Poor Repertoire; WM – Writhing Movements.

*Techniques and Models:* ASD – Acceleration Standard Deviation; CIMA – Computer-based Infant Movement Assessment; CPP – Cerebral Palsy Predictor; CSD – Standard Deviation of the Center of Motion; FFT – Fast Fourier Transformation; HOJD2D – Histograms of Joint Displacement 2D; HOJO2D – Histograms of Joint Orientation 2D; ICC – Intraclass Correlation Coefficient; LDA – Linear Discriminant Analysis; LDOF – Large Displacement Optical Flow; LLGMN – Log-linearised Gaussian Mixture; LR – Logistic Regression; MEMD – Multivariate Empirical Mode Decomposition; MRF – Multi-label Markov Random Field; NGBS – Naive Gaussian Bayesian Surprise; PCKh 2.0 – Percentage of Correct Keypoints in Relation to Head Segment Length (two times the head segment length); QMEAN – Quantity of Motion Mean; Qmedian – Quantity of Motion Median; QSD – Quantity of Motion Standard Deviation; RBF – Radial Basis Function Kernel; RF – Random Forests; RPSR – Robust Point Set Registration; SMIL – 3D Skinned Multi-Infant Linear (Based on SMPL Model for Adults); SMPL – Skinned Multi-Person Linear Model; SVM – Support Vector Machine; VSD – Standard Velocity Deviation.

**Table 3**

Challenges, limitations, and future directions of different computer vision sensing and data processing approaches.

	<b>Current approaches / Problems</b>	<b>Future directions / Improvements/ Challenges</b>
<b>Sensors</b>	<p><b>Current approaches</b></p> <p>Mostly 2D single cameras</p> <p><b>Problems</b></p> <p>Only 2D information</p> <p>Occlusions</p>	<p><b>Future directions</b></p> <p>Multiple 2D cameras</p> <p>3D (depth) sensors</p> <p>Pressure mat sensors</p> <p><b>Improvements</b></p> <p>3D information</p> <p>Less occlusions</p> <p>More information due to multi-sensory integration</p>
<b>Data</b>	<p><b>Current approaches</b></p> <p>Small datasets</p> <p><b>Problems</b></p> <p>Not enough data to employ deep learning methods</p> <p>Not publicly available - no benchmarking possible</p> <p>Incorrect or incomplete data in some cases, e.g., inaccurate outcome labelling due to lack of longitudinal studies, the inclusion of incorrect age-specificity cases, use of low-inter-rater agreement or small rater-group or lack of experienced raters in data labelling, disorders or gender misrepresentation</p>	<p><b>Future directions</b></p> <p>Collect more data</p> <p>Make it publicly available</p> <p>Make use of home videos</p> <p>Employ DL methods</p> <p>Make use of transfer learning (e.g., Tan et al., 2018)</p> <p><b>Challenges</b></p> <p>Need to solve anonymisation issue (automated techniques for face detection and replacement can be applied)</p> <p>Development of methods which can cope with different light conditions, resolution, frame rate</p>
<b>Body areas of interest</b>	<p><b>Current approaches</b></p> <p>Mostly movement of arms, legs, head</p> <p><b>Problems</b></p> <p>Incomplete information of full-body movement</p>	<p><b>Future directions</b></p> <p>Hand, fingers, feet</p> <p>Eye movement data</p> <p>Mimic</p> <p><b>Challenges</b></p> <p>Integration and analysis of multimodal information</p>
<b>Motion tracking</b>	<p><b>Current approaches</b></p> <p>Mostly in 2D space</p> <p><b>Problems</b></p> <p>Only 2D information</p>	<p><b>Future directions</b></p> <p>Full-body tracking in 3D using well-established methods in DL (e.g., DeepLabCut and OpenPose frameworks)</p> <p><b>Challenges</b></p> <p>DL methods need to be adapted to infants</p>
<b>Motion encoding</b>	<p><b>Current approaches</b></p> <p>Conventional features based on: displacement, distance, velocity, acceleration, speed, and time</p> <p><b>Problems</b></p>	<p><b>Future directions</b></p> <p>Motion encoding using well-established methods from robotics:</p> <p>Dynamic Movement Primitives (e.g., Ijspeert, Nakanishi, Hoffmann, Pastor, &amp; Schaal, 2013), Gaussian Mixture Models (e.g., Calinon, 2016; Khansari Zadeh &amp; Billard, 2011), Probabilistic</p>

	<b>Current approaches / Problems</b>	<b>Future directions / Improvements/ Challenges</b>
	<p>Only 2D features</p>	<p>Movement Primitives (Paraschos, Daniel, Peters, &amp; Neumann, 2018)</p> <p>Learn features from expert knowledge during observation (e.g., Silva et al., 2018, 2019)</p> <p><b>Improvements</b></p> <p>3D features</p> <p>New motion encoding and features</p>
<p><b>Classification algorithms</b></p>	<p><b>Current approaches</b></p> <p>Conventional ML methods, e.g., SVM, Decision Trees, Neural Networks, Hidden Markov Models</p> <p>Supervised learning without feedback during learning</p>	<p><b>Future directions</b></p> <p>Employ ANN, DL</p> <p>Employ Interactive Machine Learning (learning with feedback)</p> <p><b>Improvements</b></p> <p>Better models with more accurate predictions</p> <p><b>Challenges</b></p> <p>More data is needed</p>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript