

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Using DeepLabCut to recognize early motor development patterns associated with Neurodevelopmental Disorders

Caruso Angela¹, Marzena Oliveira Ribas (previously Marzena Szkodo)¹, Martina Micai¹, Massimo Bernava², Gennaro Tartarisco³, David López Pérez⁴, Maria Fazio⁵, Przemyslaw Tomalski⁴, Maria Luisa Scattoni¹

¹Coordination and Promotion of Research, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy

²Institute for Chemical-Physical Processes, National Research Council of Italy, Viale Ferdinando Stagno d'Alcontres 37, 98158 Messina, Italy

³Institute for Biomedical Research and Innovation, National Research Council of Italy, 98164 Messina, Italy

⁴Institute of Psychology, Polish Academy of Sciences, 00-378 Warsaw, Poland

⁵Department of Mathematics, Computer Science, Physics and Earth Sciences, University of Messina, Viale F. Stagno d'Alcontres 31, 98166 Messina, Italy

Corresponding author: Martina Micai (e-mail: martina.micai@iss.it).

This work was supported by READS project (MIMIT, Prog. n. F/180026/04/X43), MSCA-ITN-2018 - European Training Networks under Grant agreement ID: 814302 (SAPIENS) "Shaping the social brain through early interactions", by the Italian Ministry of Health Project (project code: PNRR-MAD-2022-12376472) "BABY@NET: A technology-based National Surveillance Network for the early identification of Autism Spectrum Disorder and other Neurodevelopmental Disorders in at-risk populations, and by the grant from the National Science Centre, Poland (2019/32/C/H6/00199),

ABSTRACT Early identification of Neurodevelopmental Disorders (NDD) allows for faster intervention, which in turn improves clinical outcomes and reduces the individual and societal costs associated with the diagnosis. The aims of the study were to 1) investigate the use of the DeepLabCut (DLC) toolbox to automatically analyze the motor patterns of infants at Low Risk (LR) and High Risk (HR) for Autism Spectrum Disorder (ASD); and 2) define the critical time window in which atypical motor patterns discriminate between typically developing infants and those diagnosed with ASD or NDD. The DLC toolbox was used to train a model capable of tracking the movements of both LR and HR infants longitudinally at the ages of 10 days, 6 weeks, 12 weeks, 18 weeks, and 24 weeks. 226 videos of 87 infants (45 females), collected within the Italian Network for Early Detection of Autism Spectrum Disorder (NIDA), were analyzed. Using the Percentage of Correct Key-points (PCKh) accuracy metric, the DLC's tracking performance was verified by comparing the obtained 2D hands and feet coordinates with those extracted by the Movidet software. Furthermore, motor features were computed and fed to three classifiers: Fine Tree, RUSBoosted Trees, and Narrow Neural Network to investigate their usefulness in terms of early NDD prediction. Satisfactory PCKh results were obtained for both hands and feet (left foot: 96.6%, right foot: 96.2 %, left hand: 80.9%, right hand: 82.8%). The best classification results were obtained with the RUSBoosted classifier at the ages of 10 days and 6 weeks. The 5-fold cross-validation accuracy was 81.4%, with a true negative rate of 80.0% and true positive rate 87.5%. Our data confirm the usefulness of DLC as a low-cost approach to track infant movements during the writhing period. Early motor behavior at the ages of 10 days and 6 weeks carries valuable information that has the potential to be suitable in predicting the diagnosis of NDD.

INDEX TERMS Autism, DeepLabCut, High Risk infants, early behavior, movement tracking, Neurodevelopmental Disorders

I. INTRODUCTION

Neurodevelopmental Disorders (NDD), including Autism Spectrum Disorder (ASD), are early onset conditions, characterized by various deficits in one's personal, academic, social, or occupational functioning [1].

Early identification of NDD is critical but challenging due to a long prodromal period through mid- to late infancy [2, 3]. Thanks to the high degree of neuroplasticity in the first years of human life [4], early intervention has the potential not only to maximize individual outcomes and improve prognosis [5], but also to reduce the societal and family costs associated with the diagnosis [6]. Recently, there has been an increased interest in the early detection of NDD, including ASD, which is currently diagnosed, in the best-case scenario, when a child is around 2-3 years of age [7, 8, 9, 10]. According to various studies, differences between children later diagnosed with ASD and those typically developing (TD) ones are already present in the first year of life [11, 12]. Movement atypicality is one of the first signs that might precede social or language abnormalities in ASD [13]. Therefore, even though atypical motor development is not a core symptom of ASD, it is often examined in the context of risk assessment [14]. Indeed, early motor development has been extensively explored in high-risk infants of developing ASD, i.e., siblings of children with a diagnosis of ASD, in whom the prevalence of ASD is higher than that observed in the general population [15].

Although motor problems are among the most important co-occurring conditions in ASD [16], an autism-specific atypical motor profile has not yet been defined. Both qualitative and quantitative motor atypicalities may occur in autistic individuals. Early motor markers of ASD include postural asymmetries [17] and poor postural control [18]. Teitelbaum and colleagues [19] retrospectively analyzed videos of infants aged 4-6 months who were later diagnosed with ASD and found difficulties with lying, righting, sitting, crawling, and walking. Furthermore, some studies have shown that from the age of 12 months, autistic individuals received lower gross and fine motor scores compared to their undiagnosed peers [20-23]. Additionally, Phagava and co-authors [23] revealed differences between infants with and without ASD in general movements (GMs). GMs are present from early fetal life and are usually assessed until 20 weeks post-term, which is when intentional and antigravity movements appear and begin to dominate [24]. GMs are a part of the spontaneous movement repertoire and are easy to observe due to their frequent occurrence and long duration that allowing for their proper assessment [25]. At term age and during the first two months of life, the GMs are called writhing [26]. They are characterized by an ellipsoid form, and low to moderate speed and amplitude. Normal GMs are perceived as fluid, elegant, and complex, including rotations along the axis of the limbs. They involve the whole body with a

variable sequence of trunk, neck, arm, and leg movements. By 6-9 weeks after birth, fidgety movements (FMs) begin to emerge and slowly replace writhing. An awake and alert infant expresses them continuously. They are characterized by small, circular movements of the limbs, neck, and trunk. FMs are small in amplitude and moderate in speed. Even though, according to Einspieler [27], the fidgety GMs might still occur in infants until around 6 months of age, due to the emergence of voluntary, goal-directed movements between 15 and 20 weeks of age, the evaluation of FMs after 15 weeks post-term might be difficult and not fully reliable [24]. According to the study by Phagava and colleagues [23], autistic individuals more often presented a poor repertoire during the writhing period and abnormal or absent FMs. Some authors emphasize that infants later diagnosed with ASD may achieve motor milestones at the same time as their TD peers but might perform them in a qualitatively abnormal manner [28, 29]. Therefore, both qualitative and quantitative assessment of individual behaviors are needed, rather than global measures of motor milestone achievement.

Motor behaviors are often studied by marking people with physical, reflective markers, which are not only intrusive, but whose number and location have to be determined a priori [30], potentially influencing the natural behavior of the subject being studied. Therefore, markerless pose estimation, using computer vision, is becoming increasingly popular in the field of motion analysis. Noticeable improvements have been triggered by advances in convolutional networks [31]. In 2021, Desmarais and colleagues [31] showed that, among the leading human pose estimation methods, the most accurate techniques used various architectures, such as 3D human body models, learnable triangulation, or temporal convolutional networks, and a consensus on the best approach has not yet been reached. According to the authors [31], a well-recognized interdisciplinary pose estimation framework is DeepLabCut (DLC) [30, 32] – which aims to achieve human-like tracking accuracy using Deep Neural Networks. DLC is an open-source toolbox contained within a Python package. It is based on transfer learning with deep neural networks. Thanks to the ability to take a network trained on one task with a large, supervised dataset (in this case ImageNet), and use it for another task with a small, supervised dataset, DLC can use limited training data and accurately track user-defined features [32]. The major advantages of this approach are its powerful generalization ability and flexibility – the labels are personalized, and the user can decide which key points to track.

Recently, a software package called Movidia [33, 34] has been developed for the automatic analysis of movement of infants at risk for NDD. It allows the operator to track the infant's end-effectors in free moving conditions and automatically extract various motor features from a given

video. Movidia collects measurable and quantifiable information not based on visual scoring of an infant's motor performance, completed by clinicians or trained operators during the well-child visits.

In this study, we aim to 1) investigate the use of DLC to automatically analyze the motor patterns of infants at low and high risk for ASD; and 2) define the critical time window in which atypical motor patterns discriminate between typically developing infants and those diagnosed with ASD or NDD.

II. MATERIALS AND METHODS

A. PARTICIPANTS

The Italian Network for Early Detection of Autism Spectrum Disorder (NIDA Network) is the largest development surveillance program for infants at risk of NDD in Italy, coordinated by the National Institute of Health (Istituto Superiore di Sanità, ISS). Infants were recruited between 2012 and 2020 in pediatric hospitals and clinical research centers throughout the entire Italian territory. The study protocol was approved by the Ethics Committee of the ISS (Approval Number: Pre 469/2016). Written informed consent were obtained from parents/guardians before video recording. The inclusion criteria for infants were 1) birth weight ≥ 2500 g; 2) absence of known medical, genetic, or neurological conditions associated with ASD; 3) gestational age ≥ 36 weeks; 4) absence of major complications in pregnancy and/or delivery likely to affect brain development, and 5) Apgar index > 7 at the 5th minute.

A total of 226 videos of 87 infants (45 females), recorded longitudinally at the age of 10 days, 6 weeks, 12 weeks, 18 weeks, and 24 weeks, were analyzed. Participants were divided into groups with a high (HR, $n=50$) or low (LR, $n=37$) risk of ASD. Individuals from the HR group had an older autistic sibling, whereas those from the LR had no family history of autism. At the age of 24 or 36 months, participants' clinical outcomes were assessed by blinded expert clinicians from the NIDA Network using standardized tools/tests and structured interviews with parents for checking the presence/absence of an ASD or NDD diagnosis. The NIDA Network's comprehensive clinical protocol allows the characterization of a child's developmental profile in all domains, including motor, communication/language and social domains [35]. 19 participants received an NDD diagnosis, out of which 18 belonged to the HR group. In our sample, NDD diagnosis included ASD, Communication Disorders, Attention Deficit Hyperactivity Disorder, and Motor Disorder.

A total of 126 videos of the HR group, and 100 of the LR group were collected. 180 videos came from infants who did not receive a diagnosis during the assessment stage of the study, whereas 46 came from individuals diagnosed with NDD.

B. VIDEO RECORDING AND PREPARATION

The recordings took place at participants' homes to avoid any infants' and/or parents' discomfort. The child was lying on a green blanket. The camera was placed above the infant, at chest height. The recording lasted at least 5 minutes and aimed to acquire images of spontaneous movement of the child's full body [33]. The researchers (either a psychologist, neurobiologist, or therapist) placed the camera 50 cm away from the child, recorded the infant in a well-lit room, and did not interact with the infant during the process of data collection. Parents were invited to leave the room to avoid attracting the infant's attention which could disturb the expression of movement and to ensure a video recording of an infant spontaneously moving, based on Prechtl's General Movement Assessment [18].

The preliminary evaluation of the recordings showed that high-quality video without any interferences did not last longer than 3 minutes [33]. Therefore, 3-minute video segments where the infant was in supine position, in a condition of well-being, without crying episodes or accidental movements of the camera were selected. If the video lasted for longer than 3 minutes, then the first 3 minutes of the high-quality frames were chosen. If the 3-minute video segments did not reach high quality, they were not included in the analysis.

C. DLC MODEL TRAINING

To train the DLC model, the workflow used in Nath and colleagues [32] was followed. To track the infant's movements, we labeled the central points on the back of the hands and feet (Fig.1).

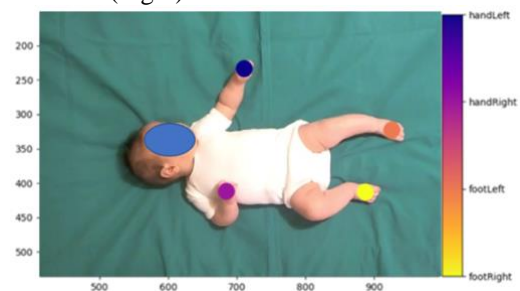


FIGURE 1. The labeling Graphical User Interface (GUI) in DLC and the tracked body parts.

To provide good generalization, the final model was trained using 3,597 labeled frames across 129 different videos. Recordings of infants from each age group were used. The algorithm was trained for 850,000 iterations using a laptop with the NVIDIA GPU (RTX 3060). The network's performance was measured as the mean average Euclidean error between the labels predicted by DLC and the manual ones. To help exclude occluded data, DLC returns not only the x and y coordinates of the body part of interest but also their probability. The user can then determine the likelihood threshold of data points. To choose the best value, the percentage of available data, as well as the train and test errors, were checked for various likelihood thresholds (from 0% to 90%, by every 10 percentage points). Exceeding the likelihood threshold of

10% did not provide a major drop in errors obtained. Therefore, to preserve as much data as possible, the likelihood threshold of 10% was chosen. The obtained train error was then 3.16 pixels, and the test error was equal to 3.47 pixels, less than 1% related to video resolution (640 x 480).

We performed a stratified 5-fold cross-validation to ensure that each fold is representative of the entire dataset in terms of class distribution. This process was carried out as follows: the dataset was randomly divided into five equal parts, maintaining the proportion of HR and LR videos in each fold. For each fold, four parts were used for training (80%), and one part was used for testing (20%). This procedure was repeated five times, ensuring that each part was used as a test set exactly once. To prevent overfitting, we ensured that the training and testing sets were completely independent for each fold. This means that no video used for training in a given fold was included in the testing set for that fold.

D. TRAJECTORIES' PROCESSING

The trained DLC model was then utilized to perform tracking of the children's hands and feet in all videos within the dataset. To get rid of outliers that could result from any tracking errors, the data were filtered in two ways. First, by using a 1-dimensional median filter with a window size equal to 7 [36]. Second, by calculating a z-score of each data point compared to the entire trajectory of the point in the respective video and eliminating those data points that were further away than 3 SDs from the mean. If the z-score of one of the two coordinates exceeded this threshold, the whole point was eliminated. After these steps, the percentages of missing data were calculated for each tracked body part and for each age group (Fig. 2). Infants tended to point the back of their hands down to the floor during the recording resulting in a high percentage of missing hand data (Fig. 2). Therefore, the computation of motor features was conducted using only the feet trajectories.

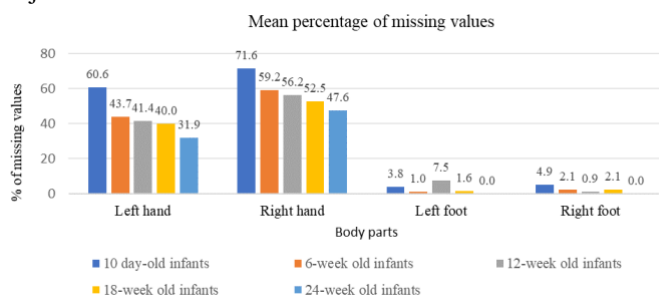


FIGURE 2. Mean percentage of missing data points for left/right hands/feet by age group.

E. TRACKING DLC PERFORMANCE VS MOVIEA

To further verify the performance of our model, we compared the hands and feet coordinates obtained with DLC with those extracted by Moveida software [33]. Moveida is a semi-automatic software designed for

tracking the end-effectors of children in a video. It requires an operator to preselect a set of parameters, including the headline, the central line of the infant's body, and the central point of the end effector. The Percentage of Correct Key-points (PCKh) accuracy metric was calculated using 50% of the head length as a distance threshold. This metric showed how often the predicted key point and the true joint were within the chosen distance limit. The Moveida-extracted trajectories were treated as the ground truth. Unlike the computation of motor features that were based on trajectories and would have been influenced by a high percentage of missing data, here the tracking comparison, not only for feet but also for hands, was done using only those data points that were available after the initial signal filtering and before any data imputation was conducted. When the DLC detected the back of the hand in a video, it was possible to check if the provided coordinates were compatible with those obtained by Moveida.

F. MOTOR FEATURES' COMPUTATION

The motor features of interest were computed in MATLAB R2022b. The missing feet data were first imputed using the default method of the `inpaint_nans` MATLAB function (Copyright (c) 2009, John D'Errico). Since the home setting did not always allow for the camera to be placed exactly 50 cm away from the baby, using pixels as a measure unit would not allow for comparison of some of the motor features between videos. For instance, mean velocities in pixels/second have different interpretations for various recordings. To overcome this problem, we used head length (measured in pixels, from chin to hairline) to normalize the data and allow comparison between subjects and time.

The list of computed features over a 3-minute period included: mean velocities and mean accelerations of the left and right foot, cross-correlation between feet velocities, cross-correlation between feet accelerations, skewness of speed distribution of each foot, periodicity in feet's trajectories, periodicity in feet's velocities, as well as the area differing from the moving average, and the area out of the SD of the moving average. These features are meaningful for the analysis of atypical motion patterns [37].

The magnitude of velocity, also referred to as speed was computed as the Euclidean distance of the central point on top of the foot between two subsequent frames, multiplied by the number of frames per second (fps) recorded in a specific video. In the dataset employed for this study, some of the videos have a frame rate of 25 fps, while others have 30 fps. By multiplying the Euclidean distance between two consecutive frames by the frame rate, we accounted for the variability in frame rates across the input videos.

Acceleration was then calculated as the difference between two subsequent speed values. Then, the mean speed and mean acceleration were computed.

The cross-correlation (CC) is a measure of synchronicity of the limbs' movement [33]. The CC with zero lag was

calculated for velocity magnitude (and analogously for acceleration) between the left and right foot (Eq. (1)).

$$CC_{v1v2} = \frac{\sigma_{v1v2}}{\sqrt{\sigma_{v1}^2 \cdot \sigma_{v2}^2}} \quad (1)$$

where σ_{v1v2} is the covariance of \mathbf{v}_1 and \mathbf{v}_2 , σ_{v1}^2 is the variance of \mathbf{v}_1 , and σ_{v2}^2 is a variance of \mathbf{v}_2 .

The skewness (v) was calculated to evaluate the speed distribution (Eq. (2)):

$$v = \frac{\frac{1}{n-1} \sum_{i=1}^{n-1} (v_i - \bar{v})^3}{\sigma_v^3} \quad (2)$$

where n is the number of recorded frames,

$$\bar{v} = \frac{1}{n-1} \sum_{i=1}^{n-1} v_i$$

$$\sigma_v = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (v_i - \bar{v})^2}$$

To determine the periodicity parameter (Eq. (3)), the trajectory (or velocity magnitude) signal of each video was divided into three equal parts [37]. An arithmetic mean was calculated for each of these parts. Then, the signal values that intersected with the mean were calculated. Further, the mean distance between the intersections and their SD were calculated. Since the videos in the dataset were recorded with various numbers of frames per second, the distance between the intersections was calculated in seconds, rather than in number of frames.

Periodicity:

$$P_{l,n} = \frac{1}{\sigma_{l,n} + \bar{d}_{l,n}} \quad (3)$$

$$P_{feet} = \sum_n P_{leftfoot,n} + \sum_n P_{rightfoot,n}$$

where $P_{l,n}$ is the periodicity, $n=[x,y]$, and $l=[left_foot, right_foot]$, $\sigma_{l,n}$ is the SD and $\bar{d}_{l,n}$ is the mean distance between all consecutive intersections and P_{feet} is the merged periodicity parameter for the left and right foot.

The area differing from the moving average was calculated using the windowing width k , that corresponded to averaging over 2 s [37]. Therefore, a script was written, which first checked how many frames per second were recorded in each video, then doubles this value and rounds it up. Next, depending on whether the resulting number was even (Eq. (4) and Eq. (5)) or odd (Eq. (6) and Eq. (7)), the appropriate equations were used.

$$\bar{x}_i = \frac{1}{k} \sum_{j=i-\frac{k}{2}}^{i+\frac{k}{2}} x_j \quad (4)$$

$$A_{diff} = \sum_{i=\frac{k}{2}+1}^{l-\frac{k}{2}} |x_i - \bar{x}_i| \quad (5)$$

$$\bar{x}_i = \frac{1}{k} \sum_{j=i-\frac{k-1}{2}}^{i+\frac{k-1}{2}} x_j \quad (6)$$

$$A_{diff} = \sum_{i=\frac{k+1}{2}}^{l-\frac{k-1}{2}} |x_i - \bar{x}_i| \quad (7)$$

where \bar{x}_i is the moving average of the i -th frame, k is the window width, x_j is the detected position in the x direction

in the j -th frame, A_{diff} is the area differing from the moving average, and l is the number of frames of the video. The same equations were used for the movement in y direction.

Since the value of the area differing from moving average depends on the length of a specific video, the parameter was normalized (Eq. (8)).

$$A_{norm} = \frac{A_{diff}}{l-k} \quad (8)$$

where A_{norm} is the normalized value of the area differing from the moving average.

A merged parameter for both feet (A_{feet}), which adds up the calculated areas of both spatial axes to one parameter was also calculated (Eq. (9)).

$$A_{feet} = \sum_n A_{norm,leftfoot,n} + \sum_n A_{norm,rightfoot,n} \quad (9)$$

where $n=[x,y]$.

As there is always some deviation of the trajectory from the moving average, this parameter tends to be always greater than 0 [37]. The calculation of the area out of SD of the moving average provided information about higher deviations from a smooth movement.

G. NDD RISK CLASSIFICATION AND MOTOR ASSESSMENT BEST AGE

Except for the above-described motor features, gender and the NDD risk status (HR/LR) of each infant were added to the dataset. Then, three models were trained using the MATLAB Classification Learner App: Fine Tree, RUSBoosted Trees, and Narrow Neural Network.

To determine the best age for the assessment of motor behavior, the classifiers were trained separately for each age group but also, given a different movement characteristic in different stage of development, for infant groups divided based on the GM period: writhing (10 days and 6 weeks) and fidgety (12 weeks, 18 weeks, and 24 weeks).

III. RESULTS

A. TRACKING DLC VS MOVIEA PERFORMANCE

The comparison of the DLC-derived coordinates with the ones extracted by Movidea (Fig. 3) showed that the average feet tracking accuracy (left foot: 96.6%, right foot: 96.2%) was higher than the average hand tracking accuracy (left hand: 80.9%, right hand: 82.8%), which might be the result of more feet data available for the model training. The achieved PCKh results are satisfactory and indicate a good performance of the method investigated.

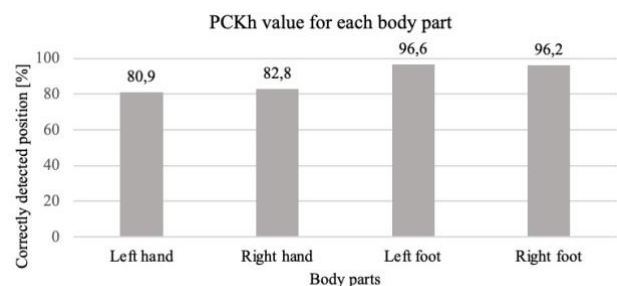


FIGURE 3. The PCKh value for each body part when comparing Movidea and DLC coordinates.

RUSBoosted Trees	FM	59.3	60.0 (66/110)	56.7 (17/30)
Narrow Neural Network	W	75.6	84.3 (59/70)	37.5 (6/16)
	FM	75.7	84.5 (93/110)	43.3 (13/30)

W: writhing period at 10 days and 6 weeks. FM: fidgety period at 12 weeks, 18 weeks, and 24 weeks. TNR: True Negative Rate. TPR: True Positive Rate. (c/a): number of correctly identified individuals/number of all subjects. Bold font indicates the results obtained by the best model.

B. CLASSIFICATION RESULTS

The classification results are presented in Table I for each age group (10 days, 6 weeks, 12 weeks, 18 weeks, and 24 weeks), and Table II for each GM period (writhing: 10 days and 6 weeks, FMs: 12 weeks, 18 weeks, and 24 weeks). As described in the methods section, the collected data were asymmetric in terms of the diagnostic status of the participants (there were more TD infants than those with NDD). As shown in Tables I and II, the RUSBoosted Trees classifier handled this better than the Fine Tree and the Narrow Neural Network, achieving more balanced results in terms of percentages of correctly identified cases and correctly detected controls. For all three classifiers, the accuracy achieved was higher for 10-day or 6-week-old infant videos than for the older infants. The best results were obtained when data from those two youngest groups were analyzed together. The 5-fold cross-validation accuracy was then 81.4%, with the true negative rate reaching 80.0% and the true positive rate equal to 87.5 % (Table II).

TABLE I
NDD VS TD CLASSIFICATION

Model	Age	Accuracy	TNR	TPR
		%	%(c/a)	%(c/a)
Fine Tree	10d	78.1	89.3 (25/28)	0 (0/4)
	6w	64.8	64.3 (27/42)	66.7 (8/12)
	12w	69.5	78.7 (37/47)	33.3 (4/12)
	18w	73.3	86.6 (33/38)	0 (0/7)
	24w	58.3	72.0 (18/25)	27.3 (3/11)
RUSBoosted Trees	10d	65.6	64.3 (18/28)	75.0 (3/4)
	6w	66.7	66.7 (28/42)	66.7 (8/12)
	12w	59.3	61.7 (29/47)	50.0 (6/12)
	18w	57.8	60.5 (23/38)	42.9 (3/7)
	24w	58.3	64.0 (16/25)	45.5 (5/11)
Narrow Neural Network	10d	78.1	89.3 (25/28)	0 (0/4)
	6w	70.4	81.0 (34/42)	33.3 (4/12)
	12w	74.6	85.1 (40/47)	33.3 (4/12)
	18w	68.9	78.9 (30/38)	14.3 (1/7)
	24w	66.7	72 (18/25)	54.5 (6/11)

TNR: True Negative Rate. TPR: True Positive Rate. d: days, w: weeks. (c/a): number of correctly identified individuals/number of all subjects. Bold font indicates the best results for each classifier.

TABLE II
NDD VS TD CLASSIFICATION DURING THE WRITHING AND FIDGETY PERIOD

Model	GMs	Accuracy	TNR	TPR
		%	%(c/a)	%(c/a)
Fine Tree	W	75.6	82.9 (58/70)	43.8 (7/16)
	FM	63.6	77.3 (85/110)	13.3 (4/30)
	W	81.4	80.0 (56/70)	87.5 (14/16)

DISCUSSION

The aim of this study was to investigate the usefulness of DLC for the analysis of infant motor patterns in high-risk infants for NDD, taking a step towards the development of an early screening tool for the NDD detection. To evaluate the DLC tracking performance, the mean average Euclidean error between the labels predicted by DLC and the manual ones was verified, and the data obtained by DLC were compared with the coordinates extracted from the same videos using the semi-automatic Movidea software [33]. The investigated approach provided reliable tracking of the infants' hands and feet in the analyzed videos.

Our data are in line with previous studies using similar approaches with the aim of developing automated standardized methods for the quantitative analysis of spontaneous movements in infants at high-risk (i.e., siblings and preterm infants) for NDD. The results showed that complexity indices of infants' hand and foot movements may be potential candidates for detecting developmental outcomes in high-risk infants [38, 39, 40]. A variety of bodily movement features at 4 months of age may be used as predictors in classifying infants with low and high autistic-like behaviors [41].

In the current study, the trajectory-based motor features were computed and fed to Fine Tree, RUSBoosted Trees, and Narrow Neural Network classifiers. Even though only the feet data were complete enough to be used for the analysis, promising results were obtained. The highest 5-fold cross-validation accuracy of 81.4% was obtained when analyzing and combining the data of writhing period (10 days and 6 weeks of age) using the RUSBoosted Trees. The true positive and true negative rates were 87.5% and 80.0%, respectively. The RUSBoost algorithm was specifically designed to improve the performance of models trained on skewed data. It applies the RUS technique, which randomly removes examples from the overrepresented class [42] and, despite its simplicity, has been proven to be very effective [43]. This classifier provided the most balanced results in terms of correct identification of cases and controls. The current work suggests that the RUSBoosted Trees classifier trained on the two early groups of infants (10 days and 6 weeks after birth) could be included in a system for the early detection of NDD.

This study analyzed a large number of videos collected early during the infants' development within the NIDA Network across the entire Italian territory. Moreover, the dataset included five age groups. Thanks to the NIDA video collections, we were able not only to train a classifier that is capable of distinguishing between TD individuals and those diagnosed with NDD at the age of 24/36 months, but also to

determine the valuable time point in development to perform an early motor analysis. Our analysis shows that the motor features computed at the age of 10 days and 6 weeks are useful in terms of early identification of NDD and might provide a better distinction between TD and NDD infants than the same motor features analyzed at later stages of development (12, 18, and 24 weeks). The intentional movements observed later in the development may interfere with the assessment of quantitative analysis of motor behaviors. Even though the older age might be non-optimal for early screening using the presented approach, analysis of FMs may provide valuable information on infant motor development. Nevertheless, it seems that data recorded early in life are more suitable to be used for an automated screening tool, as they do not include intentional movements that occur later in life.

A notable limitation of this study is the focus on extracting parameters related to the magnitude of movement only, without considering its direction. Addressing this aspect represents a significant avenue for future research. In the future, other key points and body parts could be considered for tracking with DLC to find out whether additional information could further improve the classification results. Moreover, in this study the selection of 3-minute segments, described in the methods, was done manually, which is a time-consuming task. For a fully automated tool that could be used in real-world settings, software that would automatically choose only the appropriate video parts should be developed.

This study provides an example of the application of DLC, which holds significant promise for the early assessment of neurodevelopmental delays in research and clinical settings, in combination with gold standard tools. Infant motor development can be objectively quantified and can predict neurodevelopmental outcomes. Future research should aim to promote the development of automated tools that allow to detect potential neurodevelopmental deficits early enough to provide timely intervention and improve clinical outcomes. Automated tools may be extremely useful in clinical settings where human and economic resources are often scarce.

CONCLUSION

Although more data should be collected and software for automatic video pre-processing should be developed to apply the presented approach in real-world settings, the results are very promising. The present study confirms the usefulness of DLC as low-cost approach to track infant movements in the writhing period. It has shown that the analysis of early motor patterns can predict the diagnosis of NDD, including ASD, with high accuracy. Furthermore, the very early stage of life (10 days and 6 weeks after birth) seems to be the most suitable time for using the DLC approach.

REFERENCES

[1] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM-5. American psychiatric association Washington, DC, 2013.

- [2] C. Einspieler, J. Sigafos, K. D. Bartl-Pokorny, R. Landa, P. B. Marschik and S. Bölte, "Highlighting the first 5 months of life: general movements in infants later diagnosed with autism spectrum disorder or Rett syndrome", *Res Autism Spectr Disord*, vol. 8, no. 3, pp. 286–291, Mar. 2014.
- [3] N. Deconinck, M. Soncarrieu and B. Dan, "Toward better recognition of early predictors for autism spectrum disorders", *Pediatr Neurol*, vol. 49, no. 4, pp. 225–31, Oct 2013.
- [4] F. Y. Ismail, A. Fatemi and M. V. Johnston, "Cerebral plasticity: Windows of opportunity in the developing brain", *Eur J Paediatr Neurol*, vol. 21, no. 1, pp. 23–48, Jan. 2017.
- [5] C. Lord, T. Charman, A. Havdahl, et al, "The Lancet Commission on the future of care and clinical research in autism", *Lancet*, vol. 399, pp. 271–334, Jun. 2022.
- [6] M.L. Ganz, "The lifetime distribution of the incremental societal costs of autism", *Arch Pediatr Adolesc Med*, vol. 161, no. 4, pp: 343–9, Apr. 2007.
- [7] S. L. Hyman, S. E. Levy, S. M. Myers, Council on Children With Disabilities, Section on Developmental and Behavioral Pediatrics, "Identification, evaluation, and management of children with autism spectrum disorder", *Pediatrics*, vol. 145, no. 1, Jan. 2020.
- [8] C. Okoye, C. M. Obialo-Ibeawuchi, O. A. Obajeun, S. Sarwar, C. Tawfik, M. S. Waleed, A. U. Wasim, I. Mohamoud, A. Y. Afolayan, R. N. Mbazue, "Early Diagnosis of Autism Spectrum Disorder: A Review and Analysis of the Risks and Benefits", *Cureus*, vol. 15, no. 8, Aug. 2023.
- [9] M. Hadders-Algra, "Early Diagnostics and Early Intervention in Neurodevelopmental Disorders-Age-Dependent Challenges and Opportunities", *J Clin Med.*, vol. 10, no. 4, Feb. 2021.
- [10] M. Micai, F. Fulceri, A. Caruso, A. Guzzetta, L. Gila, M. L. Scattoni, "Early behavioral markers for neurodevelopmental disorders in the first 3 years of life: An overview of systematic reviews", *Neurosci Biobehav Rev*, vol. 116, pp: 183–201, Sep. 2020.
- [11] G. Purpura, V. Costanzo, N. Chericoni, M. Puopolo, M. L. Scattoni, F. Muratori, F. Apicella, "Bilateral Patterns of Repetitive Movements in 6- to 12-Month-Old Infants with Autism Spectrum Disorders", *Front Psychol*, vol 8, Jul. 2017.
- [12] S. J. Sheinkopf, J. M. Iverson, M. L. Rinaldi and B. M. Lester, "Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder", *Autism Res*, vol. 5, no. 5, pp: 331–339, Aug. 2012.
- [13] J.M. Iverson, "Developing language in a developing body, revisited: The cascading effects of motor development on the acquisition of language", *Wiley Interdiscip Rev Cogn Sci*, vol. 13, no. 6, Nov. 2022.
- [14] L. Zwaigenbaum, S. Bryson and N. Garon, "Early identification of autism spectrum disorders", *Behav Brain Res*, vol. 251, pp. 133–146, Aug. 2013.
- [15] S. Ozonoff, G. S. Young, A. Carter, D. Messinger, N. Yirmiya, L. Zwaigenbaum, S. Bryson, L. J. Carver, J. N. Constantino, K. Dobkins, T. Hutman, J. M. Iverson, R. Landa, S. J. Rogers, M. Sigman, and W. L. Stone, "Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study", *Pediatrics*, vol. 128, no. 3, pp. e488–e495, Sep. 2011.
- [16] M. Micai, L. M. Fatta, L. Gila, A. Caruso, T. Salvitti, F. Fulceri, A. Ciaramella, R. D'Amico, C. Del Giovane, M. Bertelli, G. Romano, H. J. Schünemann, M. L. Scattoni, "Prevalence of co-occurring conditions in children and adults with autism spectrum disorder: A systematic review and meta-analysis", *Neurosci Biobehav Rev*, vol. 155, Dec. 2023.
- [17] G. Esposito, P. Venuti, S. Maestro and F. Muratori, "An exploration of symmetry in early autism spectrum disorders: Analysis of lying", *Brain Dev*, vol. 31, no. 2, pp: 131–138, Feb. 2009.
- [18] J. E. Flanagan, R. Landa, A. Bhat and M. Bauman, "Head lag in infants at risk for autism: a preliminary study", *Am J Occup Ther*, vol. 66, no. 5, pp. 577–585, Sep-Oct. 2012.
- [19] P. Teitelbaum, O. Teitelbaum, J. Nye, J. Fryman and R. G. Maurer, "Movement analysis in infancy may be useful for early diagnosis of autism", *Proc Natl Acad Sci USA*, vol. 95, no. 23, pp.13982–7, Nov.1998.
- [20] J. Brian, S. E. Bryson, N. Garon, W. Roberts, I.M. Smith, P. Szatmari and L. Zwaigenbaum, "Clinical assessment of autism in high-risk 18-month-olds". *Autism*, vol. 12, no. 5, pp. 433–456, Sep. 2008.
- [21] R. E. Landa and E. Garrett-Mayer, "Development in infants with autism spectrum disorders: a prospective study". *J Child Psychol Psychiatry*, vol. 47, no. 6, pp. 629–638, Jun. 2006.
- [22] R. J. Landa, A. L. Gross, E. A. Stuart and M. Bauman, "Latent class analysis of early developmental trajectory in baby siblings of children with autism", *J Child Psychol Psychiatry*, vol. 53, no. 9, pp. 986–96, Sep. 2012.

[23] H. Phagava, F. Muratori, C. Einspieler, S. Maestro, F. Apicella, A. Guzzetta, H. F. R. Prechtl, and G. Cioni, "General movements in infants with autism spectrum disorders", *Georgian Med News*, no. 156, pp: 100–105, Mar. 2008.

[24] L. Zuk, "Fetal and infant spontaneous general movements as predictors of developmental disabilities", *Dev Disabil Res Rev*, vol. 17, no. 2, pp. 93–101, 2011.

[25] C. Einspieler and H. F. R. Prechtl, "Prechtl's assessment of general movements: a diagnostic tool for the functional assessment of the young nervous system", *Ment Retard Dev Disabil Res Rev*, 11(1):61–67, 2005.

[26] B. Hopkins and H. F. R. Prechtl, "A qualitative approach to the development of movements during early infancy", *Clinics in developmental medicine*, vol. 94, pp.179–197, 1984.

[27] C. Einspieler, "Abnormal spontaneous movements in infants with repeated sleep apneas", *Early Hum Dev*, vol. 36, no. 1, pp: 31–48, Jan. 1994.

[28] R. J. Landa, A. L. Gross, E. A. Stuart and A. Faherty, "Developmental trajectories in children with and without autism spectrum disorders: the first 3 years", *Child Dev*, vol. 84, no. 2, pp: 429–442, Mar- Apr. 2013.

[29] S. Ozonoff, G.S. Young, S. Goldring, L. Greiss-Hess, A.M. Herrera, J. Steele, S. Macari, S. Hepburn and S. J. S. Rogers, "Gross motor development, movement abnormalities, and early identification of autism", *J Autism Dev Disord*, vol. 38, no. 4, pp. 644–56, Apr. 2008.

[30] A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis and M. Bethge, "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning", *Nat Neurosci*, Vol. 21, no. 9, pp.1281–1289, Sep. 2018.

[31] Y. Desmarais, D. Mottet, P. Slangen and P. Montesinos, "A review of 3D human pose estimation algorithms for markerless motion capture", *Comput Vis Image Underst*, vol. 212, pp.103275, 2021.

[32] T. Nath, A. Mathis, A. C. Chen, A. Patel, M. Bethge and M. W. Mathis, "Using DeepLabCut for 3D markerless pose estimation across species and behaviors", *Nat Protoc*, vol. 14, no. 7, pp. 2152–2176, Jun. 2019.

[33] W. Baccinelli, M. Bulgheroni, V. Simonetti, F. Fulceri, A. Caruso, L. Gila and M. L. Scattoni, "Moveida: A software package for automatic video analysis of movements in infants at risk for neurodevelopmental disorders", *Brain Sci*, vol. 10, no. 4, pp. 203, Mar. 2020.

[34] A. Caruso, L. Gila, F. Fulceri, T. Salvitti, M. Micai, W. Baccinelli, M. Bulgheroni and M. L. Scattoni. "Early motor development predicts clinical outcomes of siblings at high-risk for autism: Insight from an innovative motion-tracking technology", *Brain Sci*, vol. 10, no. 6, pp. 379, Jun 2020.

[35] A. Caruso, M. Micai, L. Gila, F. Fulceri, M. L. Scattoni. "The Italian Network for Early Detection of Autism Spectrum Disorder: Research Activities and National Policies", *Psychiatr Danub*, vol. 33(Suppl 11), pp: 65-68, Dec. 2021.

[36] D. L Pérez, Z. Laudańska, A. Radkowska, K. Babis, A. Kozioł and P. Tomalski, "Do we need expensive equipment to quantify infants' movement? A cross-validation study between computer vision methods and sensor data" in *EEE International Conference on Development and Learning (ICDL)*, Beijing, China, 2021, pp. 1–6.

[37] L. Meinecke, N. Breitbach-Faller, C. Bartz, R. Damen, G. Rau and C. Disselhorst-Klug, "Movement analysis in the early detection of newborns at risk for developing spasticity due to infantile cerebral palsy", *Hum Mov Sci*, vol. 25, no. 2, pp. 125–144, Apr. 2006.

[38] 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, W. H. Lee, "Deep learning-based quantitative analyses of spontaneous movements and their association with early neurological development in preterm infants", *Sci Rep*, vol. 24, Feb. 2022.

[39] H. Abbasi, S. R. Mollet, S. A. Williams, L. Lim, M. R. Battin, T. F. Besier, A. J. C. McMorland, "Deep-Learning Markerless Tracking of Infant General Movements using Standard Video Recordings", *Annu Int Conf IEEE Eng Med Biol Soc*, Jul. 2023.

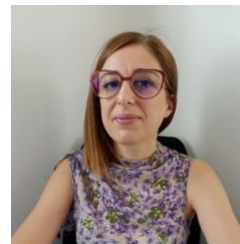
[40] M. W. Park, H. I. Shin, M. S. Bang, D. K. Kim, S. H. Shin, E. K. Kim, E.S. Lee, H. I. Shin, W. H. Lee, "Reduction in limb-movement complexity at term-equivalent age is associated with motor developmental delay in very-preterm or very-low-birth-weight infants", *Sci Rep*, vol. 14, no. 1, Apr. 2024.

[41] H. Doi, N. Iijima, A. Furui, Z. Soh, R. Yonei, K. Shinohara, M. Iriguchi, K. Shimatani and T. Tsuji, "Prediction of autistic tendencies at 18 months of age via markerless video analysis of spontaneous body

movements in 4-month-old infants", *Sci Rep*, vol. 12, no. 1, pp. 1–12. Oct. 2022.

[42] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse and A. Napolitano, "RUSBoost: A hybrid approach to alleviating class imbalance", *IEEE Trans. Syst. Man Cybern. A Syst. Hum*, vol. 40, no. 1, pp. 185–197, Jan. 2010.

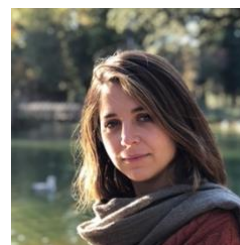
[43] J. Van Hulse, T. Khoshgoftaar and A. Napolitano, "Experimental Perspectives on Learning from Imbalanced Data", vol. 227, pp. 935–942, Jun. 2007.



C. CARUSO. Dr. Angela Caruso completed her Master of Sciences in Biology and received a PhD in Psychobiology and Psychopharmacology at the School of Behavioural Neuroscience, Sapienza University of Rome, Italy in 2017. Additionally, Dr. Caruso holds a second-level Master of Science degree in Principles and Practice of Systematic Reviews and Meta-Analysis in the Biomedical Field, also from the University of Padua, Italy. She is currently a permanent Researcher at the Italian National Institute of Health, Rome, Italy within the National Observatory for Autism. Her research focus is to examine early markers of Neurodevelopmental Disorders in high-risk infants through the implementation of novel technologies. Since 2015, Dr Caruso collected motor, social, vocal, and clinical data of infants of the Italian Network for Early Detection of Autism Spectrum Disorder (NIDA) Network devoted to identifying early markers of Neurodevelopmental Disorders. She has extensive experience in designing and analyzing data from novel software for infants' movements and kinematics analysis.



M. OLIVEIRA RIBAS. Marzena Oliveira Ribas holds a Master of Science degree in Psychology with a specialization in Clinical Psychology and a Bachelor's degree in Biomedical Engineering. As an Early-Stage Researcher for the ITN Marie Curie project SAPIENS at the Istituto Superiore di Sanità in Rome, Italy, she conducted experimental research focused on the early recognition of neurodevelopmental disorders, utilizing new technologies and innovative data analysis methodologies.



M. MICAI. Dr. Martina Micai, born in Ferrara, Italy in March 1988, earned her Bachelor of Science degree in Cognitive and Psychobiological Sciences and a Master of Science degree in Neurosciences and Neuropsychological Rehabilitation from the University of Padua, Italy. During the period from 2013 to 2016, she was a Marie Skłodowska-Curie Actions Innovative Training Network (ITN) Early Career Researcher Fellow within the EU-funded project Language and Perception (LanPercept). This fellowship coincided with her pursuit of a joint Ph.D. in Psychology at the University of Seville, Spain, and in Language and Linguistics at the Norwegian University of Science and Technology in Trondheim, Norway. Additionally, Dr. Micai holds a second-level Master of Science degree in Principles and Practice of Systematic Reviews and Meta-Analysis in the Biomedical Field, also from the University of Padua, Italy. Following her academic pursuits, she assumed a postdoctoral research position at the University of Reading, United Kingdom. Since 2018, Dr. Micai has been a permanent Researcher at the National Autism Observatory of the National Institute of Health in Rome, Italy. Within this role, she contributes to the investigation of early biomarkers of Neurodevelopmental Disorders, utilizing advanced technologies within the Italian Network for Early Recognition of Autism Spectrum Disorders (NIDA). Her expertise encompasses experimental research focused on the mental health, diagnosis, and treatment of individuals with neurodevelopmental disorders and psychiatric conditions.



M. BERNAVA. Dr. Giuseppe Massimo Bernava received his degree in computer engineering from the University of Catania (Italy) in 2004, followed by a PhD in computer science from the University of Milan in 2009. His doctoral research focused on the applications of machine learning for decision support in the medical field. Currently, he is a Technologist at the Italian National Research Council (CNR) - IPCF. His research interests lie in the design and development of medical devices using innovative software tools and approaches, particularly in leveraging artificial intelligence for understanding biological phenomena. His expertise includes image and video processing for feature identification and tracking through deep learning techniques. Bernava is also the author or co-author of numerous scientific papers published in international journals.



G. TARTARISCO. Dr. Gennaro Tartarisco received a MSc in Biomedical Engineering at the University of Pisa in 2009 and a Ph.D. degree in Automatic, Robotic and Bioengineering from the same university in 2013. Dr. Tartarisco was a Research Fellow at the Institute of Clinical physiology of National Research Council of Italy, in collaboration with the Interdepartmental Research Centre "E. Piaggio" of the Faculty of Engineering of

University of Pisa from 2013 to 2015. Since 2016 He is a PhD Researcher at the Institute for Biomedical Research and Innovation (IRIB) of National Research Council of Italy, Messina unit. His research approach spans from mobile health technology and computational science in medicine. His interests lie around the recent advances of wearable healthcare systems and telemedicine, coupling with signal processing and data mining techniques such as machine learning and deep learning algorithms. One of the main contributions of his research works is related to the monitoring of health parameters, as physiological and cardiovascular clues, vital body signs and the patients' social and environmental context to understand biological phenomena and assisting medical diagnosis, rehabilitation, and early detection of diseases. He has published over 80 works in peer-reviewed international journals.



D. LÓPEZ PÉREZ. Dr David López Pérez is a dynamic systems expert, having earned a Ph.D. and a Post-doc at the Institute of Psychology, Polish Academy of Science, and the University of Warsaw, Poland. With a research focus on the application of non-linear dynamical system methods to study human and animal behavior. He investigated the role of movement in the early development of attention in infants, he specializes in non-linear methods, interactions, movement dynamics, eye-tracking, near-infrared spectroscopy, and magnetic resonance imaging. Dr. López Pérez's research interests extend to parent-infant interactions, examining how infants dynamically learn to synchronize their actions with their parents through the use of state-of-the-art movement extraction algorithms and dynamical system methods.



M. FAZIO. Dr. Maria Fazio holds her PhD degree in Advanced Technologies for Information Engineering from the University of Messina, Italy. She is Associate Professor in Computer Science at the University of Messina (Italy). In her carrier, she was involved in several national and international research projects and scientific collaborations. She is a member of the Editorial Board of several international journals and has been Guest Editor of several Special Issues published in international journals. She was chair and organizer of international conferences and workshops. Currently, she is responsible of the CINI "HPC: Key Technologies and Tools" (HPC-KTT) National Laboratory and vice-chair of the Master' Degree international course in Data Science at the University of Messina. She is the author of more than 200 papers in the field of distributed systems and computing technologies. Her

research interests are focused on Computing continuum, with particular reference to intelligent microservice orchestration, auto-configuring mesh networks at the Edge, security in IoT-Edge ecosystems. She is co-founder of Alma Digit, which is a SME and academic Spin-Off at University of Messina aimed at Implementing automation of processes in the Cloud.



P. TOMALSKI. Dr. Przemysław (Przemek) Tomalski graduated in 2005 with a MA degree in Psychology from the University of Warsaw. For his PhD he investigated subcortical face processing under the supervision of Mark H. Johnson and Gergely Csibra at Birkbeck, University of London (UK). He then studied the effects of poverty on infant neurocognitive development in multi-ethnic and multi-lingual population of East London (UK), conducting the first resting EEG study of infants from families differing in socio-economic status. Later projects (funded by the Nuffield Foundation) involved predicting language development using mobile eye-tracking tasks and testing the effectiveness of early attention training in infants from low-SES families. Between 2012 and 2019 his Babylab at the University of Warsaw, Poland, investigated the effects of SES and household chaos on early attention and dyadic social interactions, as well as neural mechanisms of audiovisual speech processing in infancy (using EEG and fNIRS neuroimaging). He coordinated the Marie Skłodowska-Curie Innovative Training Network (SAPIENS, EU Horizon2020) for training young scientists in methods for studying social interactions and functional brain development. His current research focuses on the mechanisms through which infant-parent interactions shape early attention and communicative development. This work is funded by the National Science Centre, the National Agency for Academic Exchange and the European Commission.



M.L. SCATTONI. Dr. Maria Luisa Scattoni obtained her Ph.D. in Pharmacology and Toxicology from Sapienza University of Rome in 2005. She holds the position of Research Director at the Italian National Institute of Health, Rome, Italy. Dr. Maria Luisa Scattoni is the Research Director and Coordinator of the Italian National Autism Observatory and the Italian

Network for the Early Recognition of ASD (NIDA) since 2010. In this role, Dr. Scattoni facilitates collaboration and clinical networks between paediatricians, child neuropsychiatrists, and neonatologists throughout Italy. MLS responsibilities also encompass the development of standardized protocols, as well as the creation and validation of innovative tests and technologies for early ASD recognition. Dr. Scattoni's work extends to monitoring the developmental trajectories of high-risk populations, such as siblings of children with ASD, children born preterm, and infants born small for gestational age. Additionally, since 2018, she coordinated the development of the National guidelines and organization of services for people with Autism Spectrum Disorder and their families. She has published more than 140 articles in the field. Her contributions have been recognized through various research grants and awards, including the prestigious Mention of Honor in the Research Sector as part of the 100 Italian Excellences Award in 2018.