Quantification and Visualisation of Interpersonal Synchrony using Wearable Sensors: A Case Study on Autistic and Neurotypical Children

Yanke Sun, *Member, IEEE*, Sally Day, *Member, IEEE*, Thomas J. Gilbert, Maria Bell, Antonia F de C Hamilton, Jamie A Ward

***Abstract*— Interpersonal synchrony (IS), a key indicator of social interactions, is traditionally assessed through video data and manual coding methods, a process that is time-consuming and subjective. This study presents an automated sensor-based framework for quantifying and visualizing IS using time series data collected from wearable sensors, demonstrated through a case study of interactions between autistic and neurotypical children in classroom settings. We evaluated time series similarity measures, including Cross-correlation (CC), Dynamic Time Warping (DTW), and Cross-Wavelet Analysis (XWA), as features for machine learning (ML) models that classify interaction levels, where ground truth labels are derived from video-coded motor coordination as a behavioral proxy for IS. Results show that these similarity-based features outperform conventional statistical features in distinguishing high and low IS using ensemble classifiers. We further compare two approaches for identifying pseudosynchrony: a surrogate data analysis for threshold estimation and a supervised learning approach for direct prediction, providing a systematic evaluation of their methodological trade-offs that has been largely overlooked in prior synchrony research. The developed visualization tools enable dynamic tracking of interaction patterns while filtering out pseudosynchrony. The proposed workflow offers a scalable, objective, and reproducible alternative to manual coding, addressing a key gap in the current literature and supporting broader applications in social, developmental, and rehabilitation research.**

***Index Terms*— Machine learning, Time series analysis, Wearable sensors, Human–computer interaction, Social interaction, Data visualization, Educational measurement**

This work was supported in part by the European Research Council (ERC) under Grant 899779, SocSensors Project

Yanke Sun is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (e-mail: yankesun89@gmail.com).

Sally E. Day is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (e-mail: sally.day@ucl.ac.uk).

Thomas J. Gilbert is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (e-mail: thomas.j.gilbert@ucl.ac.uk).

Maria Bell is with the Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, United Kingdom (e-mail: maria.bell@slam.nhs.uk).

Antonia Hamilton is with the Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, United Kingdom (e-mail: a.hamilton@ucl.ac.uk).

Jamie A. Ward is with the Department of Computing, Goldsmiths, University of London, London SE14 6NW, United Kingdom (e-mail: j.Ward@gold.ac.uk).

I. Introduction

THE evaluation of synchrony in social interactions has attracted multidisciplinary attention because of its role in social signal processing, computational neuroscience, developmental psychology, and child psychiatry [1]. Interpersonal synchrony (IS), a measure of behavioural coordination, the tendency of individuals to align their actions and behaviours temporally during interaction, is a crucial feature of social interactions [2].

High levels of synchrony are linked to positive outcomes such as social bonding [3], relationship intimacy [4], negotiation success [5], therapeutic efficacy [6], and enhanced classroom learning. Conversely, reduced synchrony is associated with social isolation, loneliness, and adverse mental health outcomes [7]. In particular, reduced social coordination is reported in individuals with autism [8], [9], a population that has difficulty responding to social cues and understanding people's emotions, significantly influences their ability for social interaction and communication [10].

Quantifying interpersonal synchrony has the potential not only to advance our understanding of social interaction but also to support assessment, monitoring, and hypothesis generation, particularly in neurodevelopmental conditions such as autism. However, the subtlety and complexity of synchrony pose significant challenges for measurement and analysis due to lack of analytical tools. Traditionally, IS has been studied through manual coding methods, which involve trained observers scoring video recordings of interactions based on predefined criteria [11]. While this method offers rich qualitative insights, it is time-consuming, labour-intensive, and subject to human bias and error [1].

Recent advances have led to automated approaches, which aim to provide more scalable and objective analyses. Among this, Video-based tools have become popular for data collection [12], [13], but face line-of-sight issues and may raise significant privacy concerns. Alternatively, wearable technology, particularly Inertial measurement unit sensors (IMU) sensors, have emerged as particularly promising tools for capturing movement data [14]–[16]. To analyze these signals, time series similarity measures such as Cross-correlation (CC), Dynamic Time Warping (DTW) and Cross-wavelet Analysis (XWA) are used for capturing different types of synchrony, including in-phase and out-of-phase synchrony [17]. Furthermore, the use of machine learning (ML) techniques has enhanced the accuracy and scalability of behavioral classification tasks in this context [18]. Despite these advancements, the lack of accessible, generalizable automatic tools has limited exploration in IS research [1].

This paper proposes an automated framework for quantifying and visualizing interaction dynamics using accelerometer data, specifically through a real-world dataset that includes diverse interactions between teachers, children with and without autism in the classroom. The main contributions include:

- We propose an integrated framework that combines time-series similarity and machine learning (ML) models to automatically estimate levels of interpersonal synchrony from wearable sensor data, offering a scalable alternative to manual video coding.
- To our knowledge, this is the one of the first studies to systematically compare surrogate data analysis (SDA) and ML-based methods for pseudosynchrony detection, highlighting the methodological trade-offs between statistical thresholding and data-driven prediction.
- We develop an interactive visualization suite to track synchrony dynamics across time and participant pairs, incorporating automated exclusion of pseudosynchrony for clearer interpretation of genuine social coordination.
- We present a complete, replicable pipeline, from data acquisition to feature extraction, classification, and visualization, demonstrated on real-world classroom data involving children with and without autism. This workflow offers a reusable methodological foundation for future research on interpersonal synchrony in comparable settings.

II. Related Work

A. Wearable Sensors and Automated Coding Methods for Quantifying Interaction

In combination with automated coding methods, wearable sensors offer a promising solution for conducting more efficient, objective, and scalable analysis of interaction dynamics. Time-series analyses like Cross-correlation (CC) and Dynamic Time Warping (DTW) and time-frequency spectral analyses such as Cross-wavelet Analysis (XWA), are popular automatic coding approach for studying IS.

Windowed CC analysis offers a simple approach to quantify the dynamic change in synchrony [19]. This method has been applied to physiological signals, such as heart rate and skin conductance, capturing the dynamics of social interactions between dyads [20]. However, the method's effectiveness depends on the selecting parameters such as window size, maximum lag, and window step [21].

DTW has been applied in investigating behavioural synchrony in speech and language patterns within clinical interactions [22] and to study how gestures kinematically relate to each other within ensembles during discourse [23]. XWA is applied to elbow angles to identify whether interpersonal motor coordination emerges between two participants when they intentionally do not coordinate their movements between each other [24]. It is also utilised to find the correlation between coordinate points of body joints to explore nonverbal synchrony [25]. XWA has also been applied to head accelerations between actors and audience members in theatre performance, successfully capturing interactions across various frequencies [26]. Interactions can occur when individuals move at the same frequency but in different phases [27]. This indicates that the analyses with the properties for finding similarities with a time lag to study IS is particularly advantageous.

B. Importance in Studying Synchrony in Autism

The symptoms Autism spectrum condition (ASC) often emerge during early development that ASC children are slower to grasp social signals and fail to imitate movements during interactions with people compared with neurotypical children [28]. Several studies have also found that atypical movement coordinations are shown between autistic participants in different age groups [8], [9], [29]. Early interventions, which include parent-mediated, therapist-delivered interventions and school-based strategies [30], can be critical for autistic children. Improved interpersonal coordination in autistic children has been linked to a reduction in repetitive behaviours [31] and an increase in social competence [32], [33]. Therefore, it is valuable to develop a quantitative measurement tool to quantify behavioural coordination and understand the symptoms of autism, which could help the intervention procedures. Our previous work successfully demonstrates the feasibility of identifying subtle moments of social coordination between actors and autistic children using wearable sensors from a lab-based environment [34]. This direction is further supported by the broader landscape of sensor-based ASC research, a recent systematic review highlights how AI, IoT, and wearable sensor technologies are increasingly applied to capture behavioural and motor markers of ASC across diverse real-world settings, underscoring their potential for more accurate and scalable assessment [35]. Here, we studied groups of autistic and neurotypical children in the real-world classroom study, representing a particularly challenging scenario for measurement of social interaction.

C. Using ML to Study Social Interaction

Advances in ML techniques have enabled researchers to study patterns of motor behaviour with higher accuracy [18]. Researchers have used ML models to discern modes of synchrony in hand movements, as captured by 3D depth cameras during mimicry exercises [36]. A particularly compelling application of these techniques is in autism research, where ML models have utilised features related to IS to predict autism with significant accuracy. Measures of IS derived from naturalistic video recordings have been used to classify whether individuals are autistic or neurotypical [37]. Coordination features extracted from hand and body movements during conversations have been utilised in predicting autism with an accuracy of 75.9% [38]. Comparisons among various ML models using features derived from multimodal sensors for ASD diagnosis have yielded the highest accuracy of 88.42%, demonstrating performance comparable to top-level

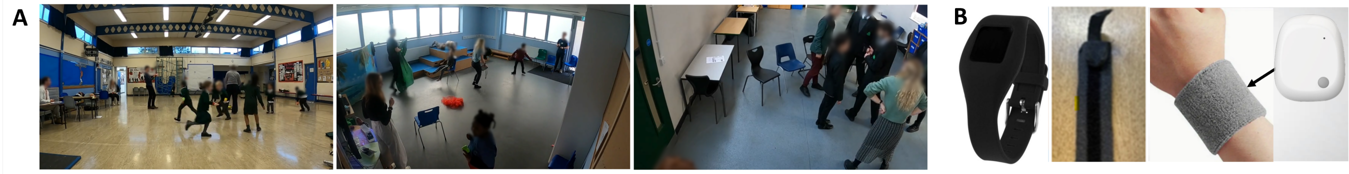

Fig. 1. Data collection contexts, wristbands and sensors, and sample data. A. The data collection contexts included a large hall in a primary school and two classrooms. B. Every participant wore a wristband (three types) with an mBient sensor which tracks wrist acceleration.

human experts [39]. To date, most studies employing ML in synchrony have focused on distinguishing between types of participants (e.g., autistic vs. non-autistic) or different forms of synchrony. The application of ML in the assessment of IS remains relatively limited [36].

D. Distinguish pseudosynchrony

A critical challenge in the study of IS, lies in determining the threshold that separates significant synchrony from coincidental or chance synchrony [1]. Researchers have developed various techniques, many of which involve establishing a baseline based on surrogate or pseudo-data to compare with real interactions. The synchrony analysis applied on the surrogate dataset successfully constructs a threshold for judging the dyad's coordination in various studies [40]–[43], which we called the surrogate data analysis (SDA) method in this paper. Bernieri and Rosenthal introduced an early approach to generate a baseline, where video images of interaction dyads were isolated and recombined in random pairings [44]. The concept of shuffling participants or or randomizing the original time series has been developed to be part of automatic computational methods. Fourier transform-based methods randomize the phase information in the data while keeping the amplitude spectrum unchanged [45], and permutation-based methods involve randomly shuffling segments of the original data to alter the time structure [46]. Moulder el at. introduces multiple permutation-based methods, including data shuffling, segment shuffling, and data sliding [47]. By setting a minimum threshold based on the analysis of pseudo-data, the real synchrony measure exceeding this threshold is considered significant [47], [48]. Using this SDA technique to avoid capturing pseudo synchrony, a positive correlation was found between eye contact and the engagement ratings of the conversation partners [49].

E. The Need for Visualisation Tools

Tools for accurately quantifying IS can be used for research on different psychiatric and developmental conditions that affect social interaction. Linear-time diagram are used to display groups as well as one-to-one personal interactions over time [50]. The intensity of pairwise interaction have been shown in the form of heatmap called interaction matrix [34]. Graph centrality measures have been used for network analysis [51], [52], which could potentially be used to identify key individuals who serve as interaction hubs within the network. However, there are currently very few tools capable of achieving this in an automated way, and the lack of efficient and accessible tools has limited the wider exploration of IS [1]. The availability of user-friendly and accessible tools can reduce the barrier to entry for researchers and practitioners interested in studying IS and has the potential to facilitate more efficient, objective, and reproducible research in this field.

III. Experiment setup

The Socsensor dataset has been collected from children and teachers engaging in classroom activities using IMU, which provides a rich context for examining behavioural coordination of both neurotypical (NT) and autistic (ASC) children.

A. Participants

Participants in this study were from 3 distinct groups of children at 3 different schools. The young ASC group (ASCy) include 4 children aged 5-6 years from special educational needs (SEN) primary schools, whose verbal skills are rated by teachers below the level expected at 24 months of age. 7 ASC children in the older ASC group (ASCo) aged 12-17 years, attending an SEN unit within a mainstream secondary school, have verbal abilities within a secondary age range. The early years' NT group have 12 children aged 4-5 years from a mainstream primary school. There are 3 teachers, 2 teaching assistants (RA), and 2 researchers involved in this study as adult participants. The Ethics was approved by the UCL Research Ethics Committee (approval on 5975/004). Written informed consent was obtained from parents or legal guardians of all participants. Video recordings from this study were used solely for the purpose of manual coordination coding and were stored on a password-protected institutional server with access restricted to authorised members of the research team.

B. Data collection

RAs arrived at the children's school prior to each data collection session to prepare the classroom. Two GoPro cameras were set up to record the room in different directions. Figure 1 A shows the recording from three sessions of different groups. The sensors used in this study are MetaMotionR from MBIENTLAB as shown in Figure 1 B, which were controlled using an iPad to start recording before they were synced. The accelerometer and gyroscope are set to record at a 25Hz sampling rate. The data is logged and saved to the onboard memory, then downloaded to an iPad after each session.

All the sensors were put in a specially designed box and dropped together on a table once before and after the session to generate kinetic events that could be used later for data synchronisation [26]. Sensors were then fixed into the wristbands, and RA helped the children wear the wristbands properly. There are three types of wristbands, as shown in Figure 1 B; wearers chose according to their preferences.

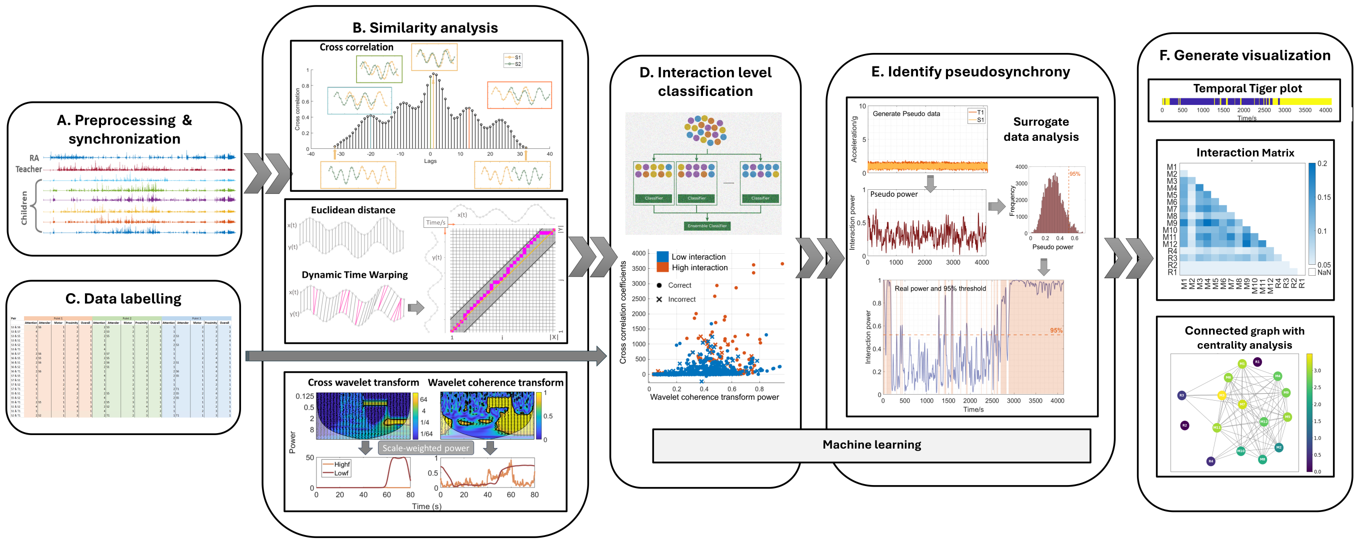

Fig. 2. Overview of the methodological pipeline. The process includes data preprocessing and synchronization, followed by interaction-level labeling. Similarity-based analyses are then applied to extract features for classifying levels of interpersonal synchrony. Subsequently, thresholds for real synchrony are identified and used to generate multiple forms of visualizations that illustrate interaction dynamics.

C. Sessions

Data were collected from ASC children during drama lessons at both SEN schools. Three teachers were present to support ASCy group, while one was present for ASCo group. Data were collected from NT children at the mainstream school as part of the after-school club. Two research scientist led the session, which incorporated age-appropriate party games. Sessions collected from 8 different days (summarized in Table I) were used in this study, which includes 3 for ASCy group, 3 for ASCo group and 2 for NT group. The sessions for ASCy incorporate music, dance, props, free play, and an interactive whiteboard. Drama games such as wink murder and charades, as well as written exercises, were included in ASCo sessions. Data were collected from NT children from part of the after-school club. The research team led the session, which incorporated age-appropriate party games, such as parachute games, Grandmother's Footsteps and Simon Says. Two RAs helped with all the sessions and were in charge of observing children and keeping records of any unintended interactions with the sensors, e.g., if children played with or removed the sensors or if children left the room. Such interruptions accounted for an average of 2% (maximum 9%) of recorded time per interrupted participant per session, confirming that data quality was not materially affected.

IV. Methods

The step-by-step framework for developing automatic tools to quantify and visualize interpersonal synchrony is illustrated in Figure 2, with each step described in detail below.

A. Data preprocessing and synchronisation

Kinetic synchronisation actions are taken before and after each session to generate an easily identifiable pattern in acceleration. Real-world synchronization errors between the sensors are corrected by aligning kinetic events in the data using interpolation technique [26]. The Euclidean norm of the 3-axis acceleration and angular rate is computed, which creates a single orientation-invariant signal $|S| = \sqrt{x^2 + y^2 + z^2}$ for each participant. High-frequency noise is then filtered out using a 4th-order Butterworth filter with 10Hz cutoff frequency. An approximately 1% temporal discrepancy was observed between the acceleration signals and the video recordings due to differences in the internal timing systems of the GoPro camera and the mBient sensors. This discrepancy was estimated by comparing the duration between two synchronized actions visible in the video with the corresponding synchronization pattern in the sensor data. To correct for this mismatch, linear scaling was applied to the sensor timestamps to align them with the video time. Following this adjustment, alignment was verified by confirming that the durations between synchronization actions, as well as the timing of other clearly identifiable activities during the sessions, were consistent across the two data streams.

B. Similarity Analysis

Windowed CC and DTW: CC measures the similarity between a signal and shifted versions of another signal. The CC is computed for all possible lags up to a maximum lag as selected, and the final CC value is chosen as the maximum correlation across all lags.

DTW measures the distance between two time series, invariant to temporal distortions [53]. It aligns signal $x(t)$ and $y(t)$ by computing the *warping path* W, a sequence of grid points $w_k = (i, j)_k$ on a $|X| \times |Y|$ grid, representing the alignment of x_i and y_j. The optimal warping path minimizes the total distance while follow some criteria set to limit the search space [54]. DTW reduces the influence of temporal shifts compared to Euclidean distance (ED) (when the warping window is set to zero) in finding similarity [55].

CC and DTW are both applied using a sliding window approach to determine pairwise similarity between sensor signals in windows. The signals are segmented into successive temporal windows, each shifted by a specified step size. The similarity values are obtained (using either DTW or CC) for each window. We choose step size of half window, since overlapping windows are used, the final similarity

is calculated by averaging successive similarity values in the overlapped area. A window size of 5 seconds was chosen based on previous studies showing that social behaviours can be reliably identified in 'thin slices' of 5 seconds [56].

Cross Wavelet Analysis (XWA): XWA includes cross-wavelet transform (XWT) and wavelet coherence transform (WCT). Firstly, continuous wavelet transform (CWT) is applied separately to two signals $x(t)$ and $y(t)$ to obtain wavelet coefficients $W_{x,\phi}(\tau, s)$ and $W_{y,\phi}(\tau, s)$. The wavelet coefficients from the two signals are then combined in two different ways: Cross wavelet transform (XWT) and Wavelet coherence transform (WCT). The XWT reveals scales with high common power [57]:

$$\boldsymbol{W_{xy}(\tau, s) = W_{x,\phi}(\tau, s) W_{y,\phi}(\tau, s)^*} \quad (1)$$

Where $*$ denotes complex conjugate. The XWT power is further defined as $|W_{xy(\tau,s)}|$. WCT finds common frequencies regardless of power in two signals, by normalizing the XWT power according to the signals' individual power [57]:

$$\boldsymbol{R^2_{xy}(\tau, s)} = \frac{|s^{-1}S(W_{xy}(\tau, s))|^2}{s^{-1}S(|W_{x,\phi}(\tau, s)|^2)s^{-1}S(|W^*_{y,\phi}(\tau, s)|^2)} \quad (2)$$

where S is a smoothing operator(see [58] for details)

After obtaining the XWT/WCT power from the two signals, weights of coefficients at different scales can be selected, which is equivalent to setting the importance of the period (frequency) range of activities to be focused. In our study, all scales have equal weights, which is equivalent to calculating the average power of XWT and WCT across selected frequency range (we select scale from 0.5s to 15s in this study). This choice was made to avoid introducing assumptions about the relative importance of different frequency bands in interpersonal movement coordination. Because the interaction dynamics in the classroom setting may occur across multiple temporal scales, treating the scales equally provides a neutral representation of movement synchrony.

C. Data labelling

Manual coding methods were designed to score the video clips according to motor coordination on a 1-5 Likert scale. Motor coordination was selected as the coding criterion as it provides an observable and operationalizable behavioral proxy for IS. Since IS is widely recognised as a key indicator of social interaction but cannot be directly observed or measured, motor coordination coding from video serves as the closest available ground truth for interaction level. A subset of video clips was selected for scoring, as full video labeling is beyond our team's capacity. An equal distribution of child-child and child-adult pairs was chosen from each session, resulting in a total of 76 pairs for analysis. Pre-analysis is performed on acceleration data using three algorithms, XWT high-frequency, WCT high-frequency and low-frequency, where potential time points of three relatively high interactions and relatively low interactions are extracted from each algorithm for each pair. This selection of video for labeling aims to balance the number of clips with high and low interactions to avoid highly skewed interaction levels. Time points in which children engaged in unintended interactions with sensors, as recorded by the RAs, were excluded. Additionally, any two selected points from the same participants and algorithms were required to have at least a 2-minute gap between them. 1273 time points were ultimately selected for 76 pairs across 8 sessions.

TABLE I
SUMMARY OF SESSION DETAILS AND LABEL DISTRIBUTION ORGANIZED BY FOLDS IN LSOCV IN THE SOCSENSOR DATASET.

Fold	Groups	Session	Dyads	Labels High	Labels Low	Labels Total
1	NT	day 1	13	76	156	232
2	NT	day 2	12	53	154	207
3	ASCy ASCo	day 1 day 1	5 10	41	229	270
4	ASCy ASCo	day 3 day 2	11 9	55	243	298
5	ASCy ASCo	day 2 day 3	7 9	38	228	266
Total count			**76**	**263**	**1010**	**1273**

Based on the 1273 time points, corresponding video clips were extracted for scoring. These video clips were then sent to 3 assistants who were blind to the algorithms and coding methods used. To assess inter-rater reliability, 10% of the clips were double-scored. Cohen's Kappa values were calculated at 0.56, 0.61, and 0.59 for raters 1&2, 2&3, and 1&3, respectively.

D. Interaction classification

The motor coordination video scores were categorized into two classes based on manually coded motor scores: 'high interaction' (Class 1) for scores exceeding two and 'low interaction' (Class 0) for scores of 1 and 2, as labels for model training. The effectiveness of two distinct feature sets was compared in predicting interaction levels, including (1) five similarity measures and (2) twelve statistical features (6 for each individual in a pair) extracted from raw data in each 5-second window. The similarity features include XWT, WCT, CC, DTW and ED. The statistical features encompass mean, standard deviation, variance, median, skewness, and kurtosis of the signals, serving as a less complex feature set for comparison.

Ensemble bagged trees were selected as the final classification model because they are well suited for datasets with heterogeneous features and relatively small sample sizes. The bagging approach reduces variance by aggregating predictions across multiple decision trees and is robust to noisy features, which is beneficial for sensor-based datasets. A preliminary comparison of model performance across several supervised learning algorithms, showed that the ensemble bagged trees achieved better performance. Therefore, this model was adopted for the final pipeline, while logistic regression (LR) was retained as a baseline model for comparison.

To completely avoid the possibility of overfitting and make use of all the data in our limited dataset, a Leave-session(s)-out cross-validation (LSOCV) approach is employed to evaluate the performance. The ground truth data is partitioned into 5 folds, with each fold containing one or two sessions and approximately the same number of samples as shown in Table I. The number of scores categorized by label types for each fold is also summarized in Table I. Performance is assessed on one fold after training on the data points from the remaining folds. This session-independent evaluation process is repeated 5 times, once for each fold. The final performance is evaluated by mean and standard deviation of the Accuracy, F1-score, and Area Under the Curve (AUC) for the Receiver Operating Characteristic (ROC) curve and Precision-Recall curve across 5 folds.

E. Identify pseudosynchrony

A critical aspect of our approach is the determination of optimal thresholds that avoiding pseudosynchrony being captured as real synchrony output. We used two approach ML and SDA and compare their performance of finding threshold according to ground truth data.

1) ML approach: Utilising the ground truth score as a reference, classification labels can be established in ML (1 for pseudo synchrony, >1 for actual synchrony) to detect real synchrony among individuals. Then, the ML model is trained based on these ground truth labels to predict real and pseudosynchrony.

2) SDA: Firstly, the pseudo acceleration data is generated based on the real acceleration data. The method for generating pseudo data are selected according to the characteristics of different similarity analyses. Then, for each pair of individuals, the similarity analyses are applied to both real and pseudo acceleration signal pairs to calculate the level of interactions before obtaining real interaction power and pseudo interaction power. The power where 95% of points in pseudo interaction power exceed is selected as the cut-off power [59]. Instances in real interaction power that exceed this 95% cutoff pseudo power are identified as real coherence. This method establishes a baseline to avoid random coordination that is identified as significant.

Generate pseudo data: Time shuffling is one of the methods used to generate pseudo data, which is suitable for algorithms like CC, DTW, and ED that are used for finding temporal patterns. Among the various time shuffling techniques, data sliding is chosen due to its computational efficiency, which involves cutting a continuous time series at a random point and repositioning the segment that follows the cut to the start of the time series [41]. A cutting point at around 60% of the data is suggested in the literature [47] and set in our study.

In contrast to CC and DTW, XWA incorporates complex phase information in the time-frequency domain. So, phase scrambling technique is introduced to generate the pseudo data. The real signal is initially transformed into the frequency domain, then phase information is shuffled randomly, while the amplitude information remains unchanged. The signal is subsequently transformed back into the time domain using this new phase information.

Combining threshold: Since different similarity analyses are introduced with different ways to generate pseudo data, combining pseudo powers obtained from different similarity analyses is only valid for algorithms employing the same pseudo-data generation methods, e.g. XWT and WCT (both using phase scrambling). The pseudo powers are combined by using the weight generated from multiple linear regression. At each time instance, the pseudo-interaction power for algorithms that need to be combined is weighted and summed to obtain a combined pseudo power before the 95% threshold is found.

F. Generate Visualizations

After the pseudosynchrony is identified, the binary output could be generated to represent the either the synchrony is real or not. By using these, we have introduced three key graphical representations: temporal tiger plots, network plots, and connected graphs with centrality analysis. The temporal tiger plots display whether people are interacting among timelines, aiming to capture the evolution of IS over time. The interaction matrices are designed to visualise group-level synchrony structures, displaying the strength of interaction for each pair on a colour scale. Lastly, the connected graphs with centrality analysis reveal the relative importance or influence of each participant within the interaction network, offering insight into the roles and engagement levels of individuals in group dynamics.

Four types of centrality mechanisms are used to analyze the interaction between people in the connected graph: degree, closeness, betweenness, and Eigenvector centrality. Degree centrality shows a node's ability to communicate directly with other nodes in the network, closeness centrality measures how near a node is to all others in the network, calculated as the reciprocal of the mean shortest path distance from that node to every other node and the betweenness centrality measures the number of shortest paths that pass through a certain node [60]. Finally, Eigenvector Centrality measures a proportional value to the sum of the score of its neighbours. Eigenvector centrality evaluates a node's influence by considering not just its connections but also the importance of its neighbours [61].

V. Result

A. Classification of Interaction Levels

We evaluated the effectiveness of different feature sets for classifying interaction levels using LSOCV strategy. As summarised in Table II, the ensemble bagged tree model trained

TABLE II
PERFORMANCE COMPARISON OF MODEL AND FEATURE SETS COMBINATIONS TO CLASSIFY INTERACTION LEVELS ACROSS LSOCV.

Model	LR + sim	EnT + stat	**EnT + sim**
Accuracy	86.3 (± 4.8)	85.8 (± 4.0)	**87.7 (± 3.8)**
Macro F1	63.6 (± 7.7)	63.0 (± 7.7)	**68.0 (± 9.0)**
ROC AUC	0.79 (± 0.08)	0.75 (± 0.08)	**0.82 (± 0.07)**
PR AUC	0.40 (± 0.18)	0.40 (± 0.22)	**0.47 (± 0.22)**

TABLE III
PERFORMANCE OF THE ENSEMBLE TREE MODEL USING SIMILARITY FEATURES, EVALUATED SEPARATELY ON NT AND ASC GROUPS.

Group	NT		ASC	
Interaction	low	high	low	high
Accuracy	84.1 (± 3.5)		90.1 (± 1.3)	
Precision	87.5 (± 0.2)	68.5 (± 20.4)	92.4 (± 1.5)	46.3 (± 11.6)
Recall	90.2 (± 2.2)	58.2 (± 9.8)	97.0 (± 1.3)	27.1 (± 16.9)
Macro F1	74.2 (± 6.0)		63.9 (± 9.0)	
ROC AUC	0.87 (± 0.03)		0.78 (± 0.06)	
PR AUC	0.67 (± 0.09)		0.33 (± 0.13)	

TABLE IV

RESULTS OF ML (IN BOLD) AND SDA (NON-BOLD) METHODS FOR DETERMINING THRESHOLDS TO PREDICT REAL SYNCHRONY ON NT DAY 1.

Feature	XWT		WCT		CC		DTW		ED		XWT & WCT		All5	
Accuracy	**70.7**	78.2	**66.8**	80.9	**67.7**	79.5	**67.7**	78.2	**65.5**	81.2	**53.5**	48.1	**78.0**	N/A
F1-score	**60.9**	67.5	**57.0**	61.6	**58.1**	69.1	**57.6**	46.7	**53.5**	48.1	**63.7**	65.4	**69.8**	N/A

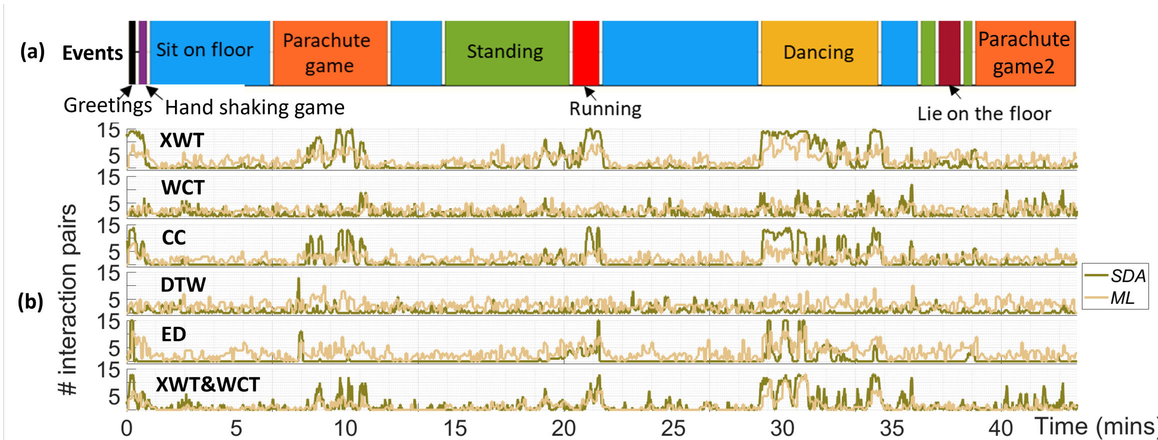

Fig. 3. Performance analysis of different threshold methods according to labelled classroom activities as the weak ground truths in a NT session. (a) displays the timeline of events, (b) shows the number of interaction pairs with R3 using either SDA or ML methods across time.

with similarity features (EnT + sim) achieved the highest performance across all evaluation metrics, outperforming both the statistical feature set using the same model ((EnT + stat) and the logistic regression model using same feature set (LR + sim). Further analysis was conducted to examine model performance across two participant groups using similarity features and ensemble tree models. Overall performance is higher in the NT group, with greater Macro F1 (74.2% vs. 63.9%), ROC AUC (0.87 vs. 0.78), and PR AUC (0.67 vs. 0.33). Classwise results highlight some disparity between NT and ASC group. In the NT group, the model achieves relatively balanced precision and recall across both interaction levels, including the minority "high interaction" class (precision: 68.5%, recall: 58.2%). In contrast, performance in the ASC group is markedly lower for the "high interaction" class (precision: 46.3%, recall: 27.1%), despite high recall for the majority "low interaction" class (97.0%). These findings suggest that, under class imbalance, the model tends to favor the majority class in the ASC group, leading to reduced sensitivity to the minority class.

B. Comparison of pseudosynchrony finding methods

By using video coding and weak ground truth data, the accuracy between the SDA and the ML methods is compared to find thresholds using a session with neurotypical children.

1) Performance comparison based on evaluation metrics: The performance of various methods is assessed using real ground truth with evaluation metrics as illustrated in Table IV. The ML method unsurprisingly shows that combining features significantly enhances performance metrics than single similarity features. Among the SDA methods, XWT, CC, and the combination of XWT & WCT exhibit relatively high performance; the overall F1 scores are competitive with those of the top-performing ML approach. SDA using fewer features, in particular XWT, WCT, CC and XWT & WCT can reach higher performance than the ML approach using the same features. The advantage of ML approach tend to improve when more features and data are available for training, while the SDA are more robust for using single or smaller numbers of features.

2) Performance comparison based on event codings: Figure 3 (b) depicts the total number of people interacting with a Researcher detected during this session with Figure 3 (a) presents the main events. Using the video data, increased physical synchrony was observed in handshake, parachute games, running, and dancing activities between the researcher and others in the classroom. Further validated the results shown in the evaluation metric, by using a single similarity measure, the ML method does not distinctly capture the variations of different synchrony levels in various activities, as the SDA.

If assessed using a combination of XWT and WCT, both ML and SDA methods achieve a more reasonable prediction than using a single similarity measure. The ML method shows clearer differences in synchrony between activities, while the SDA method successfully captures synchrony in the second parachute game, which was initially detected by WCT but not by the rest of the similarity measures, likely due to lower hand movement intensity compared to other high synchrony activities.

C. Visualisation Tools

The combination of all five features with the machine learning model achieved the highest performance, we selected this method for pseudosynchrony identification and applied it to generate visualizations for one NT and one ASC session.

1) Tiger Plot: Identify interactions over time: The interaction plot over time during one of the NT sessions between children M12 and all the other individuals in the classroom is displayed in Figure 4. The interactions in this NT session are classified into real and pseudosynchrony levels, with real interactions indicated in yellow in the tiger plots shown in Figure 4 (b). Researchers R3 and R4 lead different activities, which is evident in the pairwise interaction plot, where M12 and R4 have shown more IS during the Parachute Game, while M12 and R4 have shown more IS during the running and dancing.

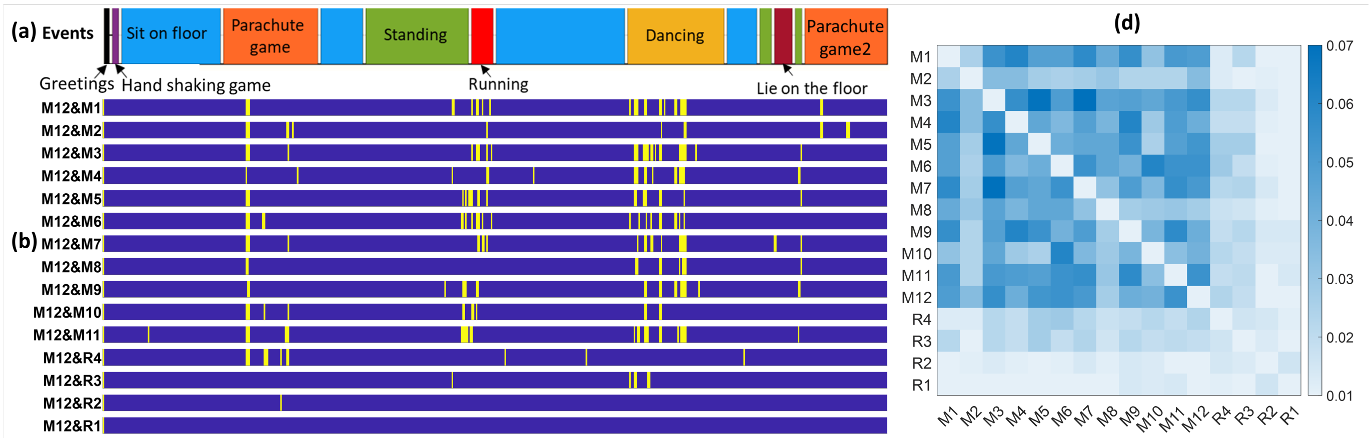

Fig. 4. Visualisation of group pairwise interactions for NT session. (a) Event coding from video. (b) Tiger plots show the interaction patterns of child M12 with all the other individuals in the classroom (d) Interaction matrix shows the percentage of interaction during the whole session for each pair.

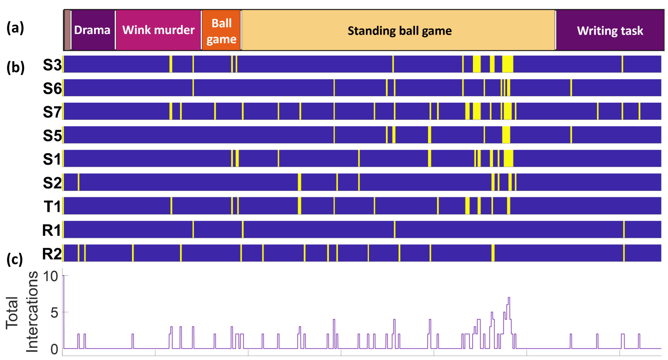

Fig. 5. Visualisation of interactions for an ASC session. (a) Event coding. (b) Tiger plot showing instances where individual interacts with other participants. (c) Total number of individuals interacting over time.

A corresponding visualisation for an ASC session is presented in Figure 5. The tiger plot in (b) represents the overall interaction of a selected individual, defined as any instance where the individual engages with another participant. As shown in (c), a higher number of interactions is observed during physically engaging activities such as the standing ball game, whereas fewer interactions occur during more structured teaching sessions (e.g., drama and writing tasks), where students are primarily working independently and interaction is largely limited to teacher-led communication.

2) Interaction Matrice: Identify pairwise interaction: The percentage of time that two individuals spend interacting in a session can be calculated, representing the proportion of time predicted as real interactions (yellow area) within Figure 4 (b). The percentage time of real interaction (PTI) is subsequently utilized to construct group interaction plots called interaction matrices, which can clearly show the dynamics of pairwise interactions during a predefined period of time as shown in Figure 4 (d). It is clear that all the children (M1-M12) interact more with children since they play games throughout the sessions. Researchers R3 and R4 lead different sections in the after-school club and show medium connections with various students. R2 and R1 show a relatively low percentage of interaction with people in the classroom since their main job is recording, and they are not involved much in the class. M2 is the least engaged in class, exhibiting low IS with almost everyone in the classroom.

3) Connected graph and graphic analysis: The Connected graph is generated with each node representing an individual and using distances between nodes to indicate the strength of pairwise PTI as shown in Figure 6. The PTI of less than 0.02 is set to 0. This visualization effectively highlights clusters of strong interaction patterns (students) and isolated individuals (R1 and R2). The lower the degree centrality value (e.g. R1-R4) have fewer direct connections within the network. In contrast, nodes with higher degrees (e.g. most of student nodes shown in light yellow) have more direct interactions, reflecting stronger local connectivity in the network. The higher value of betweenness (e.g M3) shows a higher presence of nodes in many high PTI in the network, which indicates that M3 occupies a structurally central position connecting different parts of the network. Similar pattern is shown in eigenvector centrality, M3 likely to have a prominent position within the overall interaction structure in the network as it connects to other high PTI nodes. In addition, the highest closeness centrality scores observed for M3 and M12 indicate that they are, on average, have highest PTI with all others in the network.

VI. Discussion

In the NT day 1 example, we see that using the combination of all 5 similarity features considerable improve the IS classifacation performance compared to single similarity feature. There is an evident improvement in ML model performance using LSOCV when incorporating similarity features compared to the baseline LR model and the same model using simpler statistical features, proving that similarity measures are useful features for discriminating interaction levels and combining them would make more benefits. In many real-world applications, as seen in our study, different facets of coordination exist in real interactions. Therefore, an integration of different similarity algorithms will better capture the various movement coordination patterns in interactions.

We also compared SDA and ML techniques to establish thresholds that distinguish real synchrony from chance

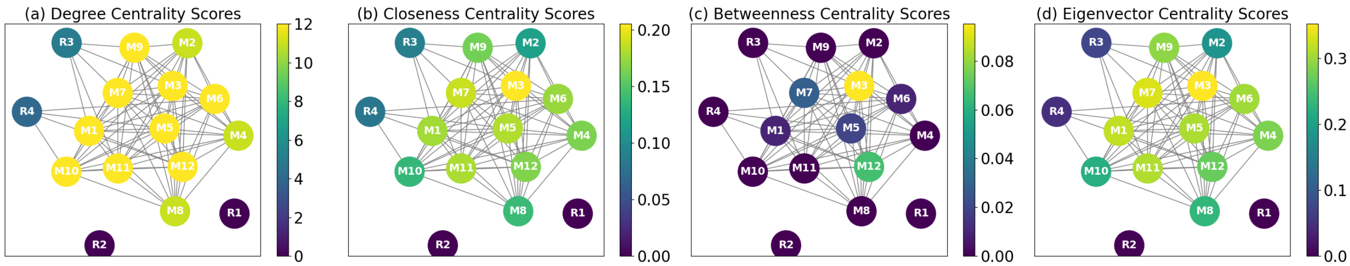

Fig. 6. Connected graph with network analysis in an after- school session for NT children using (a) Degree centrality, (b) closeness centrality (c) betweenness centrality and (d) Eigenvector centrality measures.

synchrony. With clearly labelled data for real and pseudo-synchrony, ML models can be trained to classify interactions effectively. ML is particularly useful when extensive ground truth data are available, offering a data-driven way to automate the detection of genuine synchrony. Another advantage of the ML method is its ability to combine any algorithm, unlike the SDA, which is restricted to combining algorithms using the same SDA method in this study. Although acquiring substantial and reliable ground truth data is both challenging and costly, the accuracy of performance evaluations is expected to improve as more ground truth data becomes available. In scenarios with limited or no ground truth, using SDA methods is a viable strategy for establishing thresholds. Also, it is a straightforward way to implement, though it takes time to repeat the generation of pseudo data to establish valid thresholds. One of the limitations is that the SDA methods make use of all data points in the original data to establish thresholds and some of the date points in the original data will be used for performance estimation. This introduces a risk of data leakage and could lead to an optimistic bias in the estimated performance. While transforming the original data before analysis may mitigate some of this risk, caution is warranted. It is important to note that the comparison between SDA and ML presented here is illustrative rather than definitive, and performance rankings should not be overinterpreted. This study used only a single session as an example; future work should validate both approaches across multiple independent sessions to provide a more robust assessment of their relative strengths and limitations before drawing definitive conclusions. Various visualisations have been developed based on excluding pseudosynchrony. The tiger plot can be used to monitor how interactions change over time and interaction matrix are useful for spotting interaction pairs. The network plot with centrality analysis further identify the key influence people, and people who are further away in the network. These visualisation representations are particularly useful for interaction analysis in large groups where it may be challenging for researchers and educators to discern the level of interactions visually.

However, the classification performance metrics are generally low, especially for the autistic group. One contributing factor is bias in the manually coded video labels used as ground truth. In this study, video annotations were treated as ground truth for supervised learning; however, the inter-rater reliability between coders was moderate (Cohen's Kappa values: 0.56–0.61), indicating that these labels should be considered weak ground truth rather than perfectly reliable references. Scoring the video data was challenging, even for trained raters, as the recordings captured a large classroom environment in which children frequently moved simultaneously and often occluded one another. Label uncertainty of this kind can directly affect both model training and performance evaluation. When the supervisory labels contain ambiguity or disagreement, the model may learn patterns that partially reflect annotation noise rather than true behavioural signals, which in turn limits the achievable classification performance. This challenge highlights the difficulty of relying solely on manual video coding in complex real-world environments and underscores the value of developing automated sensing approaches that complement or support human observation. In addition, the video clips used for annotation were pre-selected based on similarity measures derived from the same algorithms (e.g., XWT and WCT) that were later used as model features. While human raters assigned the final labels, this pre-selection process may introduce a sampling bias toward interaction segments that are more easily detectable by these algorithms. As a result, the reported performance may be more representative of "algorithm-detectable" interaction moments and may not fully generalise to randomly sampled or more ambiguous cases. The performance is particularly low for the classification of low interaction, which is probably due to the imbalanced nature of the data. Although we have already tried to select an equal number of high and low interaction points, as mentioned in Section IV, such imbalance reflects the real-world scenario, where interactions among autistic children are less frequent. Also, the overall sample size of the dataset are relatively small. Future investigations should consider adding more samples and more high interaction data points, although this process is labour-intensive and time-intensive. This would benifit model training and provide more robust evaluation of interaction prediction.

Thirdly, the evaluation of our session-independent cross-validation shows a relatively large difference between folds, and the performance for ASC groups is much lower than for NT groups. One factor contributing to this disparity relates to the greater heterogeneity of motor behaviour observed in autistic children. During the drama sessions, many early-year ASC participants frequently engaged in solitary actions involving large or idiosyncratic movements rather than coordinated or reciprocal interactions with peers. Such behavioural variability may reduce the consistency of movement patterns associated with interpersonal motor coordination, which the automated approach relies on for detection. In contrast, motor actions among NT children and older ASC participants were often more synchronized and socially oriented, which can produce more stable movement signatures that are easier for the algorithm to detect. Previous studies have similarly reported delays and variability in both gross and fine motor development in autistic children [62], [63], as well as atypical interpersonal movement coordination [9]. Another contributing factor is that younger autistic children are more likely to remove or manipulate wrist-worn devices and may play with wristbands more frequently [64]. These behaviours create challenges not only for data collection but also for data analysis. Interactions with the devices may introduce additional noise into the accelerometer signals and reduce the consistency of the recorded motion data, which may further affect model performance. In addition to behavioural differences, a key limitation of this study lies in the differences in data collection contexts between ASC and NT groups. The two groups were recorded under different activity settings and facilitation styles, which may influence movement dynamics and synchrony structure. Specifically, ASC data were collected during drama-based classroom activities, whereas NT data were collected during structured after-school games led by different facilitators and in different environments. These variations introduce a potential confound, as interpersonal synchrony measures are sensitive to task structure, interaction dynamics, and facilitation style. As a result, the observed differences in model performance between groups may reflect contextual and activity-related factors rather than purely in intrinsic differences associated with autism. This limits the interpretability of direct ASC–NT comparisons and reduces the extent to which conclusions can be attributed to group characteristics alone. Future work should aim to collect data under more controlled and comparable conditions across groups, or to include larger datasets that enable robust within-context validation. This would allow clearer disentanglement of activity effects from group-specific interaction patterns. In addition, we collected data from both accelerometers and gyroscopes and examined the correlation between video-coded motor scores and the similarity measures derived from both the acceleration and angular velocity magnitudes. While we found that the acceleration magnitude is a more efficient indicator, three-axis acceleration data could be leveraged as additional features to assess whether it would help with the accuracy.

Despite these limitations, this paper presents a step-by-step framework for studying IS for human interactions, from data collection to signal processing and visualization. Our findings demonstrate that using wearable sensors, automated coding methods, and visualization tools offers a promising approach to quantifying IS between groups of people, including autistic and non-autistic children. Wearable sensors present several advantages over traditional methods, such as video analysis, including cost-effectiveness and enhanced privacy. They are also highly suitable for real-world data collection, as they do not suffer from the line-of-sight issues that can arise with cameras. Therefore, wearable sensors are recognized for their excellent ecological validity and are widely accepted in almost any context. Automated coding methods, utilising different similarity measures and ML, can effectively capture the synchrony between individuals.

Automated coding methods can more efficiently and objectively analyse larger datasets, potentially enabling the discovery of subtle patterns of synchrony that might be missed by human observers. The visualization tools developed in this study provide clear visual representations of interaction dynamics at specific intervals and simplify the display of complex social relationships within large groups.

Our approach establishes a robust way to track and understand social interactions among groups. The detailed visualization figures can assist researchers and educators in identifying specific activities that promote group synchrony and understanding the interaction between different individuals. These insights can be particularly valuable in developing targeted strategies to enhance social interaction and collaborative learning within educational settings, especially beneficial for children with autism, where tailored interventions can lead to significant improvements in social integration and learning outcomes. This approach could be easily adapted to study IS across larger groups and contexts. Although our validation was limited to small cohorts of children and adults, there is potential for broader interdisciplinary applications, including activities with multiple people in the same space.

We have made the code for the tools in this study publicly available in Github[1] and the synchronized data available in OFS[2], providing researchers with accessible resources to adapt to their unique systems and study requirements easily.

VII. Conclusion

This paper presents a comprehensive framework for quantifying IS as an indicator of social interactions among diverse groups, evaluated using data collected from autistic and non-autistic children in an educational context. By extracting multiple time-series similarity measures and using them as features in machine learning models, we demonstrate the feasibility of quantifying motor synchronization from wearable sensor data. Within the scope of the present dataset, similarity-based features consistently outperformed baseline models and simpler statistical features in predicting interaction levels, and combining multiple similarity measures further improved performance. We further compared traditional SDA method with ML-based approach for identifying pseudosynchrony, which is a critical step prior to generating interpretable visualizations. Multiple forms of visualization are developed, which efficiently show the interaction dynamics at specific intervals and simplify the display of complex social relationships within large groups. Overall, the results validate the proposed framework as a scalable and automated pipeline for studying IS using wearable sensors, time-series similarity measures, and visualization tools. While the findings are based on a limited number of participants and sessions, the framework provides a reproducible methodology and practical guidelines for researchers investigating social interaction dynamics in comparable settings. Future work with larger and more diverse datasets will be essential to further assess generalisability and to explore broader applications of the approach.

[1]https://github.com/khorinaj/SocSensor-Interpersonal-Synchrony-Analysis
[2]https://osf.io/498up

ACKNOWLEDGMENT

The authors also thank the staff and pupils of Queensmill SEN School and St John Vianney Primary School for their participation and cooperation, with particular appreciation to Susan Gryphon and Elise Robinson. Additional thanks are extended to Yun Shin, Anna Fakiolas, and Nancy Beddall for their assistance with dataset scoring.

REFERENCES

[1] E. Delaherche, M. Chetouani, A. Mahdhaoui, C. Saint-Georges, S. Viaux, and D. Cohen, "Interpersonal synchrony: A survey of evaluation methods across disciplines," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 349–365, 2012.

[2] C. Bowsher-Murray, S. Gerson, E. Von dem Hagen, and C. R. Jones, "The components of interpersonal synchrony in the typical population and in autism: A conceptual analysis," *Frontiers in Psychology*, vol. 13, p. 897015, 2022.

[3] R. Mogan, R. Fischer, and J. A. Bulbulia, "To be in synchrony or not? a meta-analysis of synchrony's effects on behavior, perception, cognition and affect," *Journal of Experimental Social Psychology*, vol. 72, pp. 13–20, 2017.

[4] D. Julien, M. Brault, É. Chartrand, and J. Bégin, "Immediacy behaviours and synchrony in satisfied and dissatisfied couples." *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, vol. 32, no. 2, p. 84, 2000.

[5] A. Pentland, *Honest signals: how they shape our world*. MIT press, 2010.

[6] F. Ramseyer and W. Tschacher, "Nonverbal synchrony of head-and body-movement in psychotherapy: different signals have different associations with outcome," *Frontiers in psychology*, vol. 5, p. 979, 2014.

[7] A. Sommerlad, M. Kivimäki, E. B. Larson, S. Röhr, K. Shirai, A. Singh-Manoux, and G. Livingston, "Social participation and risk of developing dementia," *Nature Aging*, vol. 3, no. 5, pp. 532–545, 2023.

[8] A. L. Georgescu, S. Koeroglu, A. F. d. C. Hamilton, K. Vogeley, C. M. Falter-Wagner, and W. Tschacher, "Reduced nonverbal interpersonal synchrony in autism spectrum disorder independent of partner diagnosis: a motion energy study," *Molecular autism*, vol. 11, pp. 1–14, 2020.

[9] K. L. Marsh, R. W. Isenhower, M. J. Richardson, M. Helt, A. D. Verbalis, R. C. Schmidt, and D. Fein, "Autism and social disconnection in interpersonal rocking," *Frontiers in integrative neuroscience*, vol. 7, p. 4, 2013.

[10] E. M. Benssassi, J.-C. Gomez, L. E. Boyd, G. R. Hayes, and J. Ye, "Wearable assistive technologies for autism: opportunities and challenges," *IEEE Pervasive Computing*, vol. 17, no. 2, pp. 11–21, 2018.

[11] F. J. Bernieri, J. S. Reznick, and R. Rosenthal, "Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions," *Journal of Personality and Social Psychology*, vol. 54, no. 2, p. 243, 1988.

[12] Z. Hammal, J. F. Cohn, and D. S. Messinger, "Head movement dynamics during play and perturbed mother-infant interaction," *IEEE transactions on affective computing*, vol. 6, no. 4, pp. 361–370, 2015.

[13] M. E. Meza-de Luna, J. R. Terven, B. Raducanu, and J. Salas, "Assessing the influence of mirroring on the perception of professional competence using wearable technology," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 161–175, 2016.

[14] T. Choudhury and A. Pentland, "Sensing and modeling human networks using the sociometer," in *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings.* IEEE, 2003, pp. 216–222.

[15] E. Garcia-Ceja, V. Osmani, A. Maxhuni, and O. Mayora, "Detecting walking in synchrony through smartphone accelerometer and wi-fi traces," in *European Conference on Ambient Intelligence*. Springer, 2014, pp. 33–46.

[16] J. Knighten, S. McMillan, T. Chambers, and J. Payton, "Recognizing social gestures with a wrist-worn smartband," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 2015, pp. 544–549.

[17] R. C. Schmidt and M. J. Richardson, "Dynamics of interpersonal coordination," in *Coordination: Neural, behavioral and social dynamics*. Springer, 2008, pp. 281–308.

[18] S. B. Hausmann, A. M. Vargas, A. Mathis, and M. W. Mathis, "Measuring and modeling the motor system with machine learning," *Current opinion in neurobiology*, vol. 70, pp. 11–23, 2021.

[19] S. M. Boker and J. L. Rotondo, "Symmetry building and symmetry breaking in synchronized movement," *Mirror neurons and the evolution of brain and language*, vol. 42, p. 163, 2002.

[20] F. Behrens, R. Moulder, S. Boker, and M. Kret, "Quantifying physiological synchrony through windowed cross-correlation analysis: Statistical and theoretical considerations," *BioRxiv*, pp. 2020–08, 2020.

[21] D. Schoenherr, J. Paulick, S. Worrack, B. M. Strauss, J. A. Rubel, B. Schwartz, A.-K. Deisenhofer, W. Lutz, U. Stangier, and U. Altmann, "Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony," *Behavior research methods*, vol. 51, pp. 361–383, 2019.

[22] R. Lahiri, M. Nasir, M. Kumar, S. H. Kim, S. Bishop, C. Lord, and S. Narayanan, "Interpersonal synchrony across vocal and lexical modalities in interactions involving children with autism spectrum disorder," *JASA Express Letters*, vol. 2, no. 9, 2022.

[23] W. Pouw and J. A. Dixon, "Gesture networks: Introducing dynamic time warping and network analysis for the kinematic study of gesture ensembles," *Discourse Processes*, vol. 57, no. 4, pp. 301–319, 2020.

[24] J. Issartel, L. Marin, and M. Cadopi, "Unintended interpersonal co-ordination: "can we march to the beat of our own drum?"," *Neuroscience Letters*, vol. 411, no. 3, pp. 174–179, 2007.

[25] K. Fujiwara, Q. S. Bernhold, N. E. Dunbar, C. D. Otmar, and M. Hansia, "Comparing manual and automated coding methods of nonverbal synchrony," *Communication Methods and Measures*, vol. 15, no. 2, pp. 103–120, 2021.

[26] Y. Sun, D. A. Greaves, G. Orgs, A. F. de C. Hamilton, S. Day, and J. A. Ward, "Using wearable sensors to measure interpersonal synchrony in actors and audience members during a live theatre performance," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 7, no. 1, pp. 1–29, 2023.

[27] J. Issartel, T. Bardainne, P. Gaillot, and L. Marin, "The relevance of the cross-wavelet transform in the analysis of human interaction–a tutorial," *Frontiers in psychology*, vol. 5, p. 1566, 2015.

[28] T. Charman, J. Swettenham, S. Baron-Cohen, A. Cox, G. Baird, and A. Drew, "Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation." *Developmental psychology*, vol. 33, no. 5, p. 781, 1997.

[29] P. Fitzpatrick, J. A. Frazier, D. M. Cochran, T. Mitchell, C. Coleman, and e. R. Schmidt, "Impairments of social motor synchrony evident in autism spectrum disorder," *Frontiers in Psychology*, vol. 7, p. 1323, 2016.

[30] C. Lord, T. S. Brugha, T. Charman, J. Cusack, G. Dumas, T. Frazier, E. J. Jones, R. M. Jones, A. Pickles, M. W. State *et al.*, "Autism spectrum disorder," *Nature reviews Disease primers*, vol. 6, no. 1, pp. 1–23, 2020.

[31] A. Lampi, P. Fitzpatrick, V. Romero, J. Amaral, and R. Schmidt, "Understanding the influence of social and motor context on the co-occurring frequency of restricted and repetitive behaviors in autism," *Journal of Autism and Developmental Disorders*, vol. 50, pp. 1479–1496, 2020.

[32] V. Romero, P. Fitzpatrick, S. Roulier, A. Duncan, M. J. Richardson, and e. R. Schmidt, "Evidence of embodied social competence during conversation in high functioning children with autism spectrum disorder," *Plos one*, vol. 13, no. 3, p. e0193906, 2018.

[33] K. A. McNaughton and E. Redcay, "Interpersonal synchrony in autism," *Current psychiatry reports*, vol. 22, pp. 1–11, 2020.

[34] J. A. Ward, D. Richardson, G. Orgs, K. Hunter, and A. Hamilton, "Sensing interpersonal synchrony between actors and autistic children in theatre using wrist-worn accelerometers," in *Proceedings of the 2018 ACM international symposium on wearable computers*, 2018, pp. 148–155.

[35] G. Bouchouras and K. Kotis, "Integrating artificial intelligence, internet of things, and sensor-based technologies: A systematic review of methodologies in autism spectrum disorder detection," *Algorithms*, vol. 18, no. 1, 2025. [Online]. Available: https://www.mdpi.com/1999-4893/18/1/34

[36] R. Yozevitch, A. Dahan, T. Seada, D. Appel, and H. Gvirts, "Classifying interpersonal synchronization states using a data-driven approach: implications for social interaction understanding," *Scientific Reports*, vol. 13, no. 1, p. 11150, 2023.

[37] J. C. Koehler, M. S. Dong, A. M. Nelson, S. Fischer, J. Spaeth, I. S. Plank, N. Koutsouleris, and C. M. Falter-Wagner, "Machine learning classification of autism spectrum disorder based on reciprocity in naturalistic social interactions," *medRxiv*, pp. 2022–12, 2022.

[38] A. L. Georgescu, J. C. Koehler, J. Weiske, K. Vogeley, N. Koutsouleris, and C. Falter-Wagner, "Machine learning to study social interaction difficulties in asd," *Frontiers in Robotics and AI*, vol. 6, p. 132, 2019.

[39] M. Cheng, Y. Zhang, Y. Xie, Y. Pan, X. Li, W. Liu, C. Yu, D. Zhang,

Y. Xing, X. Huang *et al.*, "Computer-aided autism spectrum disorder diagnosis with behavior signal processing," *IEEE Transactions on Affective Computing*, vol. 14, no. 4, pp. 2982–3000, 2023.
[40] D. C. Richardson and R. Dale, "Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension," *Cognitive science*, vol. 29, no. 6, pp. 1045–1060, 2005.
[41] K. T. Ashenfelter, S. M. Boker, J. R. Waddell, and N. Vitanov, "Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, no. 4, p. 1072, 2009.
[42] X. Sun, A. Nijholt, K. P. Truong, and M. Pantic, "Automatic visual mimicry expression analysis in interpersonal interaction," in *CVPR 2011 WORKSHOPS*. IEEE, 2011, pp. 40–46.
[43] D. Atzil-Slonim, C. S. Soma, X. Zhang, A. Paz, and Z. E. Imel, "Facilitating dyadic synchrony in psychotherapy sessions: Systematic review and meta-analysis," *Psychotherapy Research*, vol. 33, no. 7, pp. 898–917, 2023.
[44] F. J. Bernieri and R. Rosenthal, "Interpersonal coordination: Behavior matching and interactional synchrony." 1991.
[45] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, "Testing for nonlinearity in time series: The method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 77–94, 1992.
[46] F. Ramseyer and W. Tschacher, "Nonverbal synchrony or random coincidence? how to tell the difference," *Development of Multimodal Interfaces: Active Listening and Synchrony: Second COST 2102 International Training School, Dublin, Ireland, March 23-27, 2009, Revised Selected Papers*, pp. 182–196, 2010.
[47] R. G. Moulder, S. M. Boker, F. Ramseyer, and W. Tschacher, "Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses." *Psychological methods*, vol. 23, no. 4, p. 757, 2018.
[48] Z. Kupper, F. Ramseyer, H. Hoffmann, and W. Tschacher, "Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits," *PloS one*, vol. 10, no. 12, p. e0145882, 2015.
[49] S. Wohltjen and T. Wheatley, "Eye contact marks the rise and fall of shared attention in conversation," *Proceedings of the National Academy of Sciences*, vol. 118, no. 37, p. e2106645118, 2021.
[50] M. Sathiyanarayanan and D. Pirozzi, "Linear-time diagram: A set visualisation technique for personal visualisation to understand social interactions over time," in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. IEEE, 2016, pp. 259–264.
[51] C. Correa, T. Crnovrsanin, and K.-L. Ma, "Visual reasoning about social networks using centrality sensitivity," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 1, pp. 106–120, 2010.
[52] E. C. Baek, R. Hyon, K. López, E. S. Finn, M. A. Porter, and C. Parkinson, "In-degree centrality in a social network is linked to coordinated neural activity," *Nature communications*, vol. 13, no. 1, p. 1118, 2022.
[53] A. Mueen and E. Keogh, "Extracting optimal performance from dynamic time warping," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 2129–2130.
[54] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proceedings of the 3rd international conference on knowledge discovery and data mining*, 1994, pp. 359–370.
[55] D. Sankoff and J. B. Kruskal, "Time warps, string edits, and macromolecules: the theory and practice of sequence comparison," *Reading: Addison-Wesley Publication*, 1983.
[56] N. J. Sasson, D. J. Faso, J. Nugent, S. Lovell, D. P. Kennedy, and R. B. Grossman, "Neurotypical peers are less willing to interact with those with autism based on thin slice judgments," *Scientific reports*, vol. 7, no. 1, pp. 1–10, 2017.
[57] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological society*, vol. 79, no. 1, pp. 61–78, 1998.
[58] A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," *Nonlinear processes in geophysics*, vol. 11, no. 5/6, pp. 561–566, 2004.
[59] P. Pinti, A. Devoto, I. Greenhalgh, I. Tachtsidis, P. W. Burgess, and A. F. de C Hamilton, "The role of anterior prefrontal cortex (area 10) in face-to-face deception measured with fnirs," *Social Cognitive and Affective Neuroscience*, vol. 16, no. 1-2, pp. 129–142, 2021.
[60] U. Ishfaq, H. U. Khan, and S. Iqbal, "Identifying the influential nodes in complex social networks using centrality-based approach," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 10, pp. 9376–9392, 2022.
[61] A. Alamsyah, B. Rahardjo *et al.*, "Social network analysis taxonomy based on graph representation," *arXiv preprint arXiv:2102.08888*, 2021.
[62] M. Lloyd, M. MacDonald, and C. Lord, "Motor skills of toddlers with autism spectrum disorders," *Autism*, vol. 17, no. 2, pp. 133–146, 2013.
[63] A. Mohd Nordin, J. Ismail, and N. Kamal Nor, "Motor development in children with autism spectrum disorder," *Frontiers in pediatrics*, vol. 9, p. 598276, 2021.
[64] M. Bell, E. Robinson, S. Day, T. J. Gilbert, A. F. D. C. Hamilton, and J. A. Ward, "Lessons on collecting data from autistic children using wrist-worn sensors," in *Proceedings of the 2022 ACM International Symposium on Wearable Computers*, 2022, pp. 6–10.