Identification of Infant body positions using machine learning approach

KOŁAKOWSKA Aleksandra[1], DUDA-GOŁAWSKA J.[2], LAUDAŃSKA Z.[3], ZYGIEREWICZ J.[4], TOMALSKI P.[2]

1. Faculty of Psychology, University of Warsaw, Poland
2. Neurocognitive Developmental Lab, Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland
3. Heidelberg University Hospital, Heidelberg, Germany
4. Faculty of Physics, University of Warsaw, Poland

babylab pan – Neurocognitive Development Lab

Objective

Build a research tool that automatically identifies infants' body positions from experimental data.

KEY FINDINGS

The classification model achieves high predictive accuracy, yielding **F1 scores of 97% for supine, 94% for prone, and 90% for sitting-front** positions.

Position **transition patterns** vary significantly across tasks, suggesting that infants adjust their motor strategies in response to the demands of the activity.

The **semi-continuous** classification model is being optimised on novel tasks (books, rattles) with the human-in-the-loop approach.

Supine → Prone → Sitting heels

DATA & EXPERIMENTAL SETUP

- 95 infants (4–12 months)
- 4 assessment time points
- 3 play tasks
- 12 IMU sensors (6 used for analysis)
- 3 synchronized cameras

Age \ Usable recordings	Books: 251	Rattles: 258	Manipulative: 260
4-6 months			
9-12 months			

Figure 1. The objects used during tasks: Books, Rattle shaking, and Manipulative [1].

PIPELINE

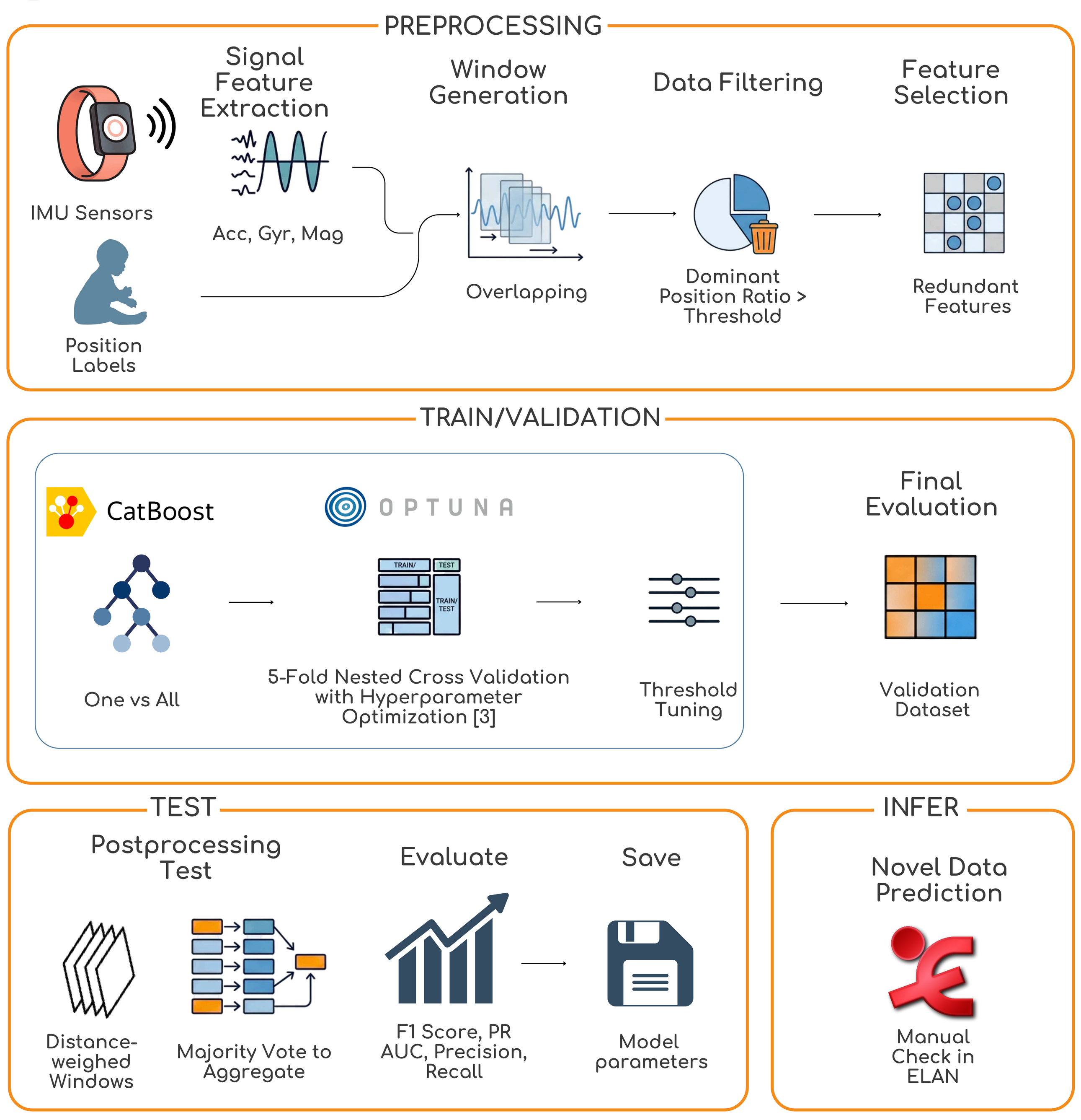

Figure 2. Comprehensive machine learning pipeline for automated position classification.

PREPROCESSING. IMU sensor data and manual ELAN [4] annotations from the Manipulative task are synchronised within labelled windows and filtered to select features.

TRAIN/VALIDATE. A custom CatBoost [2] model is optimised via nested cross-validation to manage dataset imbalance, where class representation ranges from 31% to 1%.

TEST. The model performs Semi-continuous classification generates probability scores for overlapping windows that are normalised using custom thresholds and distance-based weights.

INFER. Final position predictions are exported to ELAN format for supervised ground-truth validation by an expert.

CLASSIFICATION METRICS

5-Fold Cross Validation Confusion Matrix

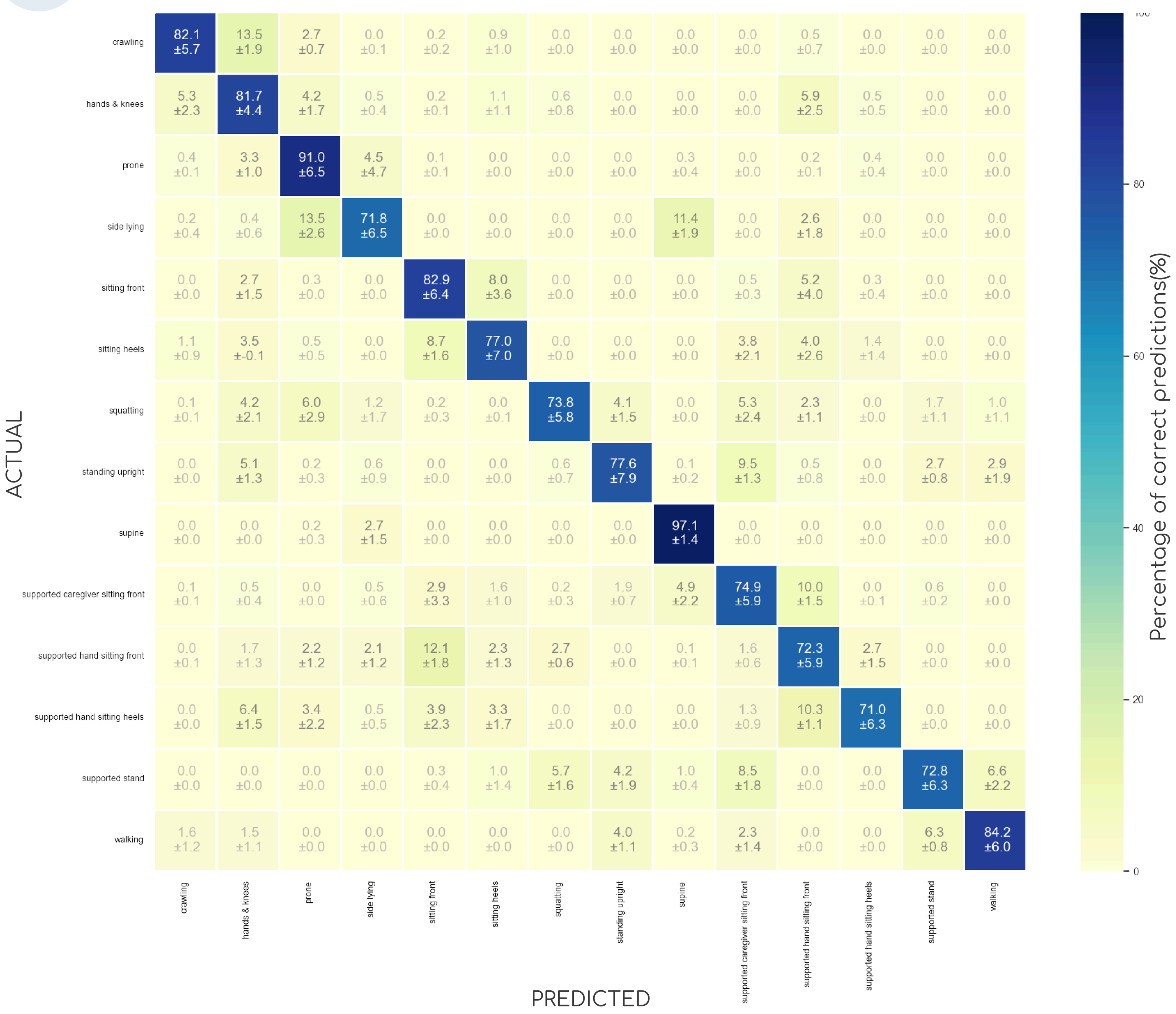

Figure 3. The results of 5-fold Cross Validation metrics on the training dataset. The data from the Manipulative dataset was stratified by the subject and split into three sets: train (70%), test (15%), and validation (15%). The group assignment was randomly shuffled 1000 times to have a sample of each position in every group. The model has the best F1 scores for supine (97%), prone (94%) and sitting front (90%)

RESULTS: REAL VS PREDICTED TIME

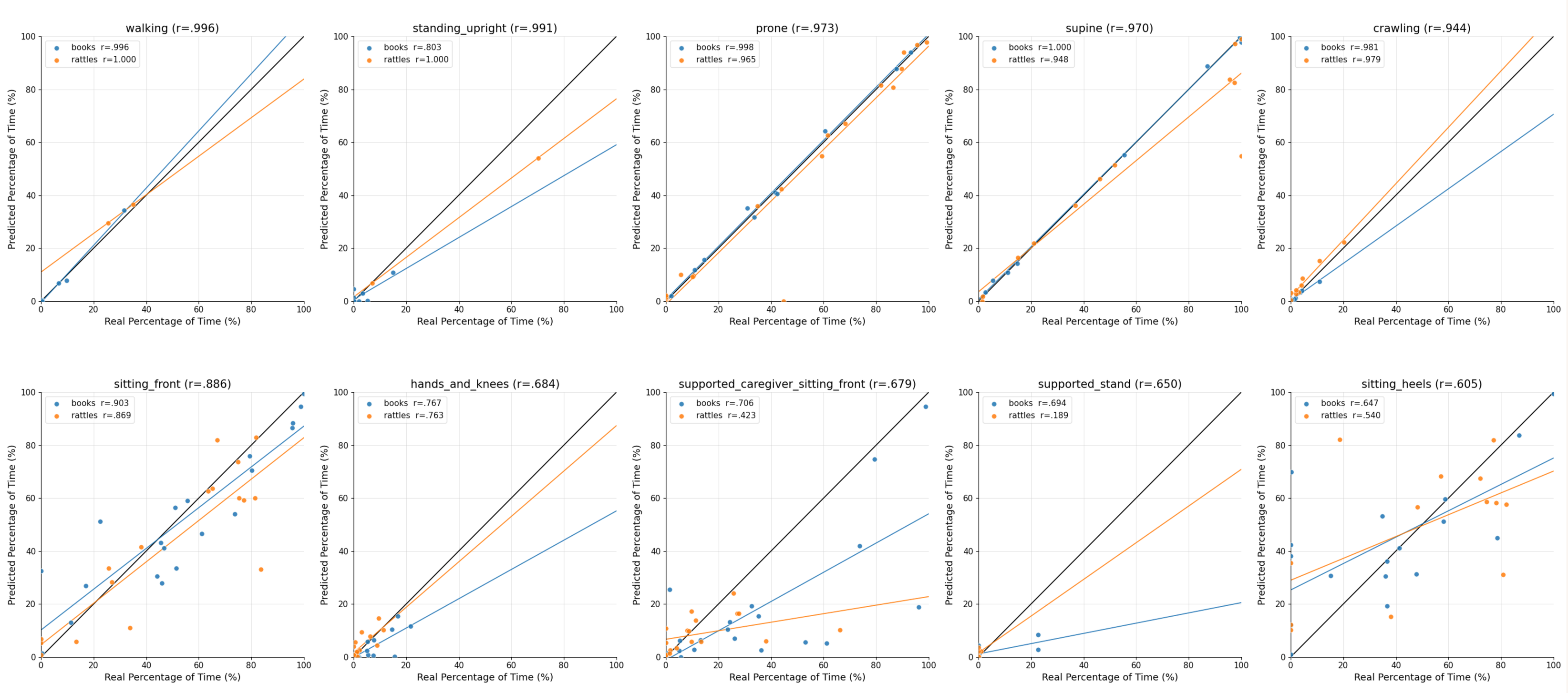

Figure 4. Correlation between actual and predicted time spent in 10 positions, using the CatBoost model. The x-axis represents the real percentage of time spent in each position, while the y-axis shows predicted values. Points closer to the line of best fit indicate more accurate predictions, with high values in positions such as walking and supine demonstrating the model's alignment with manual annotations in both books and rattles datasets. The increased variance in supported positions highlights areas of higher predictive complexity.

Sitting front → Hands & knees → Upright

ANALYSIS: POSITION TRANSITIONS

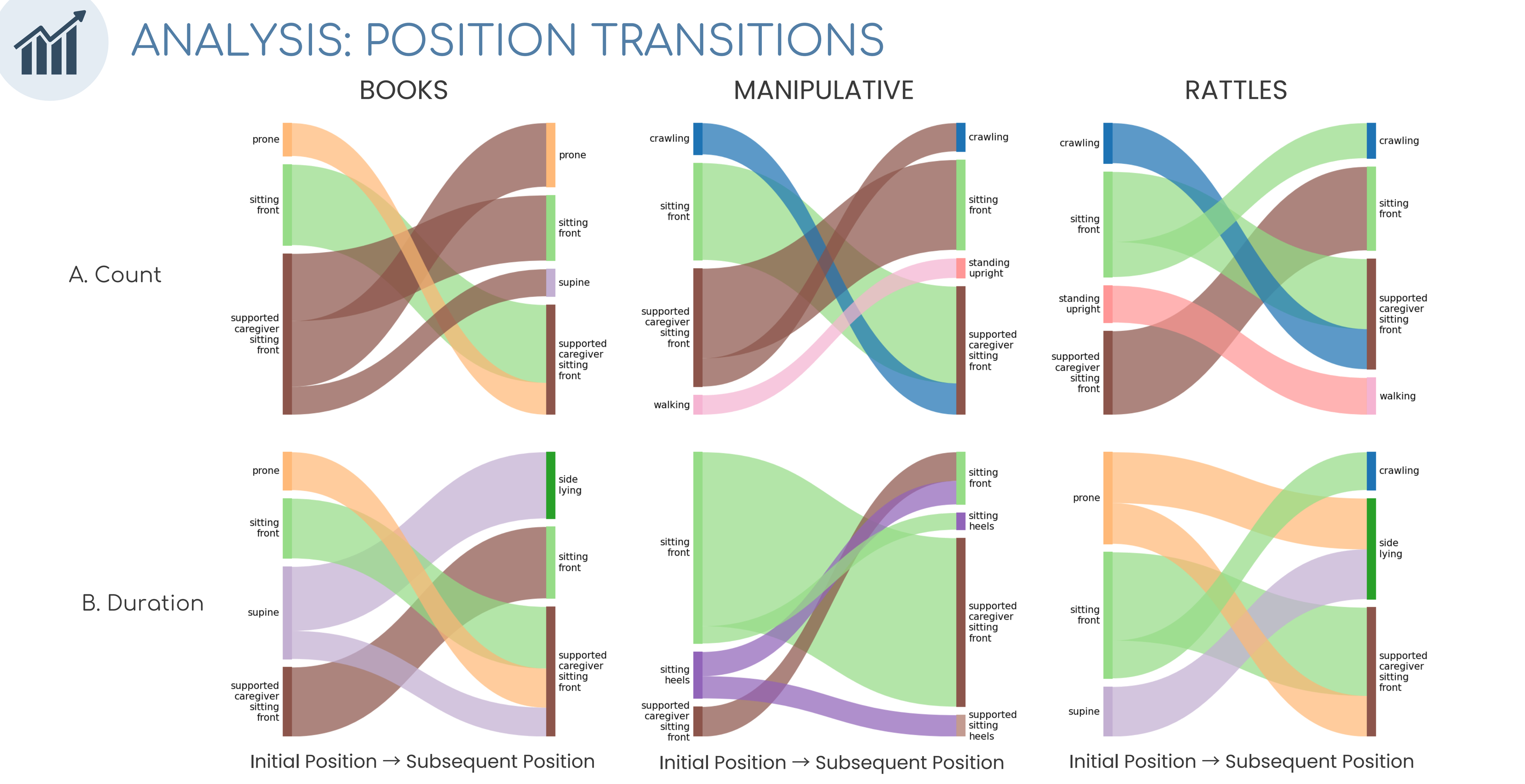

Figure 5. Comparative postural transition dynamics across tasks. Ribbon width represents the proportion of transitions, weighted by either event frequency or temporal duration. Displays show primary transitions between initial and subsequent postures, with states vertically aligned to facilitate structural comparison. Transition patterns reflect both task-specific motor exploration and methodological differences; notably, discrepancies between frequency (Fig. 5A) and duration (Fig. 5B) highlight the distinction between transient postural shifts and sustained motor engagement. While Manipulative tasks (manually annotated) show higher postural continuity, Books and Rattles (automated classification) exhibit more frequent transitions due to the nature of the sliding-window segmentation.

find us!
BABYLAB.PAN.PL

FUTURE DIRECTIONS

- Improve the model's generalizability.
- Extract infant body keypoints from video recordings using a YOLO model [5] to generate additional features.
- Compare the performance of the baseline and enhanced models.

REFERENCES

1. Duda-Goławska, J., Rogowski, A., Laudańska, Z., Żygierewicz, J., & Tomalski, P. (2024). Identifying Infant Body Position from Inertial Sensors with Machine Learning: Which Parameters Matter? Sensors, 24(23), 7809.
2. CatBoost classifier, https://catboost.ai
3. Optuna https://optuna.org/
4. ELAN transcription software, https://archive.mpi.nl/tla/elan56. Ultralytics YOLO 11 model, https://docs.ultralytics.com/models/yolo11

National Science Centre Poland
Funded by the Polish National Science Centre grants: 2018/30/E/HS6/00214 and 2022/47/B/HS6/02565 (PT).

Faculty of Physics, University of Warsaw
Institute of Psychology, Polish Academy of Sciences
Cognitive Science, Faculty of Psychology UW, Faculty of Philosophy UW

We thank all participating infants and their families.