Biomedical Signal Processing and Control 129 (2027) 111450

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Benchmarking implicit neural representations of pediatric gait waveforms against conventional biomechanical descriptors

Agnieszka Pregowska [a], Konrad Pauk [b], Jacek Tutak [c], Jolanta Pauk [d,e,*]

[a] Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland
[b] Medical University of Warsaw, Warsaw, Poland
[c] Rzeszow University of Technology, Rzeszow, Poland
[d] Bialystok University of Technology, Bialystok, Poland
[e] Glenrose Rehabilitation Hospital, Edmonton, Canada

ARTICLE INFO

Keywords:
Pediatric gait
Biomechanics
Waveform analysis
Implicit neural representations
Movement development
Gait analysis

ABSTRACT

Machine-learning representations of gait waveforms are increasingly used in movement analysis, but their added value over conventional biomechanical descriptors remains uncertain, particularly in pediatric cohorts. We evaluated whether implicit neural representations (INRs) provide useful encodings of pediatric gait waveforms for reconstruction and developmental modeling. The full cohort comprised 78 healthy children and adolescents, of whom 73 had waveform files conforming to the predefined data structure required for the repeated cross-validation analyses. Selected waveforms were modeled using several INR architectures, and reconstruction quality was assessed using R^2, RMSE, and MAE. For the INR feature-ablation analyses, one architecture was assigned to each waveform class based exclusively on reconstruction performance, independently of the age targets. Side-specific, bilateral-asymmetry, signal-specific, and descriptor-family feature sets were evaluated using ten repeats of five-fold cross-validation. A separate matched benchmark in the same 73 participants compared compact Fourier-MLP-derived features with conventional biomechanical descriptors and a fold-wise raw-waveform PCA baseline. In the INR ablation analyses, combining side-specific and bilateral-asymmetry features achieved $R^2 = 0.700 \pm 0.017$ for chronological-age regression, and Energy/AUC descriptors were the strongest individual feature family ($R^2 = 0.696 \pm 0.031$). For exploratory three-class age-group classification (4–7, 8–12, and 13–18 years), the combined INR representation reached a balanced accuracy of 0.749 ± 0.034. In the matched representation benchmark, configurations containing conventional biomechanical descriptors performed best. The biomechanics-only baseline achieved regression $R^2 = 0.837 \pm 0.070$ and balanced accuracy $= 0.798 \pm 0.109$, whereas biomechanics+PCA achieved the highest mean regression performance ($R^2 = 0.842 \pm 0.066$); this difference was not statistically significant. Adding INR-derived features did not improve regression performance and significantly reduced the evaluated classification metrics. INRs should therefore be considered complementary continuous waveform representations rather than replacements for established biomechanical descriptors.

1. Introduction

The quantitative analysis of human movement has traditionally relied on discrete spatiotemporal parameters and handcrafted biomechanical descriptors derived from kinematic and kinetic signals [1,2]. In gait analysis, measures such as joint-angle extrema, temporal events, and phase-specific summaries have been widely used to characterize developmental changes and identify pathological patterns [3–6]. Although these approaches are clinically interpretable and well established, they often require extensive preprocessing, including time normalization and cycle alignment, and may not fully capture the continuous structure of biomechanical waveforms [7,8]. These limitations have motivated the use of functional representations of gait data, including functional data analysis and principal component approaches [9,10]. Beyond gait analysis, PCA-based dimensionality reduction has also been used in other biosignal domains, including multivariate electroencephalographic analysis, to integrate correlated neurophysiological measures into compact composite representations [11]. The limitations of isolated scalar descriptors are particularly relevant in pediatric gait analysis, where developmental changes may be distributed across the entire waveform rather than localized to individual

* Corresponding author at: Bialystok University of Technology, Bialystok, Poland.
E-mail address: j.pauk@pb.edu.pl (J. Pauk).

https://doi.org/10.1016/j.bspc.2026.111450
Received 8 May 2026; Received in revised form 29 July 2026; Accepted 5 September 2026
1746-8094/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

parameters. Advances in motion-capture technology, including both marker-based and markerless systems, have further expanded the range of measurable gait signals and their clinical applicability [12].

Recent advances in machine learning have introduced implicit neural representations (INRs) as a framework for modeling continuous structured signals. Instead of representing data solely as discrete samples, INRs learn continuous coordinate-to-signal mappings, enabling representations that support interpolation and differentiation within a unified model [13–15]. INR-based methods have shown strong performance in domains such as neural rendering, 3D scene reconstruction, and continuous shape modeling [16–19]. INR-based approaches have also been investigated for structured human-body modeling and medical image reconstruction [20,21]. More broadly, they belong to a family of continuous modeling approaches that also includes neural ordinary differential equations and neural controlled differential equations [22,23].

Despite these advances, the application of continuous representation learning to biomechanical waveform analysis remains limited, and its added value over conventional biomechanical descriptors has not been systematically evaluated. Existing machine-learning approaches to gait analysis predominantly rely on handcrafted features or discretized waveform inputs [2,24,25]. Although effective for classification and regression, such approaches may not fully exploit the inherently continuous structure of gait trajectories. Recent studies have begun to compare learned latent representations with classical functional approaches, highlighting both overlaps and differences in the information captured by these representations [26]. Alternative latent-space summaries, task-supervised latent embeddings, and shared cross-subject INR embeddings were not optimized in the present benchmark; therefore, the reported results should be interpreted as an evaluation of predefined, interpretable INR-derived summaries rather than as an exhaustive assessment of all possible INR feature spaces.

A central challenge in adopting expressive neural representations such as INRs is determining whether learned features capture biologically meaningful information rather than merely achieving high reconstruction fidelity. High-capacity models can fit complex data distributions without necessarily learning interpretable or generalizable structure [27]. In gait analysis, this raises an important question: can implicit continuous representations of biomechanical waveforms encode developmentally relevant information that is useful for downstream modeling? Although continuous modeling frameworks have shown promise in other domains, their utility for capturing age-related variability in human movement remains unclear [6,28].

The present study addresses this gap by evaluating whether INR-derived representations provide biomechanically meaningful and predictively useful encodings beyond established biomechanical descriptors. Specifically, we benchmark INR-based models with respect to waveform reconstruction fidelity, chronological-age regression, and exploratory age-group classification. By comparing INR-derived and conventional handcrafted representations across multiple downstream tasks, we aim to distinguish representation fidelity from downstream developmental utility, which remains a critical consideration in applied gait analysis.

The novelty of this study therefore lies not in the proposal of a new INR architecture, but in a controlled biomechanical benchmark that distinguishes waveform reconstruction fidelity, variance preservation, and downstream developmental utility while directly comparing INR-derived features with conventional descriptors and a fold-wise raw-waveform PCA baseline.

2. Methods

2.1. Participants

The dataset comprised gait records from 78 healthy pediatric and adolescent participants (36 females and 42 males), aged 4–18 years (mean ± SD: 10.8 ± 3.9 years). Mean body height and mass were 143.2 ± 21.6 cm and 39.5 ± 15.8 kg, respectively. For descriptive purposes, participants were grouped into four age categories: 4–6 years (n = 16), 7–9 years (n = 21), 10–12 years (n = 21), and 13–18 years (n = 20). These four age categories were used only to describe the age distribution of the full recruited cohort and were not used as outcome classes in the downstream classification analyses. All participants walked independently without assistive devices and were screened to confirm the absence of musculoskeletal, neurological, or orthopedic conditions affecting gait. Exclusion criteria included recent injuries, neurological disorders, and lower-limb structural abnormalities. Anthropometric variables were recorded for each participant. Each participant completed eight walking trials at a self-selected comfortable speed. Trials were visually inspected and automatically screened for signal quality; trials affected by marker loss, force-plate targeting errors, or inconsistent gait patterns were excluded. Representative waveforms were obtained by averaging all valid trials. Left- and right-limb waveforms were analyzed separately. A further five participants were excluded from the repeated cross-validation analyses because their waveform files did not conform to the predefined data structure required by the analytical pipeline. The resulting complete-case analytical cohort comprised 73 participants. All repeated cross-validation analyses reported below were conducted in this cohort. Ethical approval was granted by the Glenrose Rehabilitation Hospital Health Research Ethics Board (HREB approval number: B-220504). Data collection was conducted at the Syncrude Centre for Motion and Balance in accordance with the Declaration of Helsinki. Written informed consent was obtained from the participants and/or their legal guardians, as appropriate.

2.2. Gait acquisition and preprocessing

Three-dimensional kinematic and kinetic data were collected using a 12-camera optoelectronic motion-capture system (Motion Analysis Corporation, Santa Rosa, CA, USA) synchronized with three multiaxial force plates (AMTI, Watertown, MA, USA). Marker trajectories were sampled at 120 Hz, whereas ground reaction forces were sampled at 1000 Hz. Reflective markers were placed on anatomical landmarks according to the laboratory's standard motion-capture protocol. Participants walked barefoot at a self-selected comfortable speed along a 10-m walkway. A static calibration trial was recorded to define participant-specific joint centers, segment coordinate systems, and anthropometric parameters. Data were processed using Visual3D software (C-Motion, Inc., Germantown, MD, USA). Marker trajectories and ground-reaction-force signals were low-pass filtered using a fourth-order, zero-phase Butterworth filter with a cut-off frequency of 6 Hz. Initial contact was identified from the vertical ground reaction force using a predefined threshold. Joint kinematics and kinetics were subsequently calculated using conventional inverse-dynamics procedures. Joint moments and powers were normalized to body mass and expressed in N m/kg and W/kg, respectively. A gait cycle was defined as the interval between two consecutive ipsilateral initial-contact events. Each gait-cycle waveform was resampled to a common normalized time base spanning 0%–100% of the gait cycle using cubic-spline interpolation. The resulting left- and right-side waveforms were retained separately for subsequent analyses [29].

2.3. Conventional biomechanical descriptors

To establish a comparative baseline, participant-level scalar feature vectors were constructed from anthropometric, spatiotemporal, and waveform-based descriptors. Anthropometric variables included body mass, height, and body mass index (BMI), together with a rescaled height parameter used to improve numerical conditioning. Spatiotemporal descriptors included walking velocity, cadence, stride and step lengths, and temporal gait-cycle characteristics such as toe-off timing

and phase durations. Symmetry indices were computed from normalized left–right inter-limb differences. Where appropriate, parameters were normalized to body height. Handcrafted descriptors were extracted from time-normalized kinematic, kinetic, and muscle-length waveforms. Global waveform features included the mean, standard deviation, minimum, maximum, quartiles, range of motion (ROM), and area under the curve (AUC). To capture phase-specific gait characteristics, each waveform was partitioned into five intervals: loading response (0%–10%), mid-stance (10%–30%), terminal stance (30%–50%), pre-swing (50%–70%), and swing (70%–100%). Within each interval, local summary features, including the mean, minimum, maximum, ROM, and AUC, were computed.

2.4. Implicit neural representation modeling and latent feature extraction

Biomechanical waveforms were modeled using implicit neural representations, in which a neural network approximates a continuous mapping from the normalized gait-cycle coordinate to waveform amplitude [13–15]. For each participant, waveform, and side, a separate INR was fitted using 101 normalized gait-cycle coordinates spanning $t \in [0,1]$.

Two reconstruction protocols were evaluated. The shallow-MLP protocol used a tanh-activated network with hidden-layer sizes of 32–8–32 units. It was fitted directly to waveform amplitudes without per-curve target standardization, using a maximum of 500 epochs and an early-stopping patience of 40 epochs.

The matched PCA–INR analysis and the associated morphology and convergence diagnostics used a compact Fourier-MLP protocol. This model used a Fourier-augmented coordinate input followed by two ReLU-activated hidden layers containing 48 and 16 units, respectively. Each target waveform was standardized individually to zero mean and unit variance before fitting. Training used a maximum of 350 epochs and an early-stopping patience of 30 epochs. Predictions were transformed back to the original waveform scale before per-signal RMSE and MAE were calculated. The aggregate convergence RMSE was calculated in the standardized domain to permit comparison across waveform classes with different physical units.

For both protocols, reconstruction metrics were evaluated at the same 101 coordinates used for model fitting. They therefore quantify in-sample functional approximation rather than reconstruction at independently held-out gait-cycle coordinates.

A broader exploratory architecture benchmark additionally considered a deep MLP, a residual MLP (ResMLP), a sinusoidal representation network (SIREN), and a higher-capacity Fourier-feature MLP. The shallow MLP in this comparison used the 32–8–32 architecture described above. The higher-capacity Fourier-feature MLP used three hidden layers containing 64 units each, a Fourier-encoding dimensionality of 16, and $\sigma = 8.0$, whereas SIREN used three hidden layers containing 64 units each and $\omega_0 = 30.0$. These exploratory architecture comparisons used different model capacities and evaluation procedures from the compact Fourier-MLP protocol used in the matched PCA–INR analysis and are therefore reported separately.

Following model fitting, the INR models were used as deterministic feature generators. Formally, for participant i, signal s, and side q, the fitted INR defines a continuous waveform approximation:

$$\hat{x}_{i,s,q}(t) = f_{\theta_{i,s,q}}(t), \qquad t \in [0,1]. \tag{1}$$

For each waveform class and side, 107 INR-derived descriptors were extracted. Descriptors were computed on the discrete grid $t_k = (k-1)/100$, $k = 1,\dots,K$ with $K = 101$. For readability the participant, signal, and side indices are omitted below. Let $x(t_k)$ denote the observed waveform and $\hat{x}(t_k)$ its INR reconstruction from Eq. (1).

The shape family contained 25 summaries of the observed and reconstructed trajectories. For $y \in \{x, \hat{x}\}$ these included amplitude, dispersion, range-of-motion, peak-timing, distributional, and extrema-related quantities,

$$\mu(y) = \frac{1}{K}\sum_{k=1}^{K} y(t_k), \qquad \sigma(y) = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\big(y(t_k) - \mu(y)\big)^2}, \tag{2}$$

$$\mathrm{ROM}(y) = \max_k y(t_k) - \min_k y(t_k), \qquad t_{\max}(y) = \arg\max_{t_k} y(t_k), \tag{3}$$

together with peak-timing, quartile, and extrema-count summaries, where $\Delta y(t_k) = y(t_{k+1}) - y(t_k)$ and $N_{\mathrm{ext}}(y) = \#\{k : \Delta y(t_k)\Delta y(t_{k+1}) < 0\}$.

The derivative family contained six descriptors. The first and second derivatives of the reconstruction, $\hat{x}'(t)$ and $\hat{x}''(t)$, are available directly from the fitted INR. The family comprised their means and standard deviations, $\mu(\hat{x}')$, $\sigma(\hat{x}')$, $\mu(\hat{x}'')$, $\sigma(\hat{x}'')$, together with the first- and second-derivative reconstruction errors

$$\mathrm{RMSE}_d = \sqrt{\frac{1}{K_d}\sum_{k}\Big(\hat{x}^{(d)}(t_k) - \Delta^{(d)}x(t_k)\Big)^2}, \qquad d = 1, 2, \tag{4}$$

where $\Delta^{(d)}$ denotes the dth order finite difference of the observed waveform and $K_d = K - d$.

The energy/AUC family contained 12 descriptors. For $y \in \{x, \hat{x}, |x - \hat{x}|\}$ the signed area, absolute area, signal energy, and curvature energy were computed by trapezoidal integration,

$$A_{\mathrm{s}}(y) = \int_0^1 y\,dt, \quad A_{\mathrm{a}}(y) = \int_0^1 |y|\,dt, \quad E(y) = \int_0^1 y^2\,dt,$$
$$C(y) = \int_0^1 (y'')^2\,dt, \tag{5}$$

yielding $3 \times 4 = 12$ descriptors.

The latent family contained 52 descriptors and was used in the signal-specific INR feature-ablation pipeline. Let $h_{i,s,q}(t) = (h_1(t), \dots, h_C(t))^\top$, with $C = 12$, denote the 12-dimensional latent trajectory used by the feature-extraction procedure. Here, $C = 12$ refers to the fixed latent representation used for descriptor construction and should not be interpreted as the width of the final architecture-specific hidden layer. For each channel $c = 1,\dots,C$ four summaries were calculated,

$$\mu(h_c), \qquad \sigma(h_c), \qquad \mathrm{range}(h_c) = \max_k h_c(t_k) - \min_k h_c(t_k),$$
$$\overline{\Delta}(h_c) = \frac{1}{K-1}\sum_{k=1}^{K-1}\Delta h_c(t_k), \tag{6}$$

giving $4C = 48$ channel-wise descriptors. Four additional global summaries described the overall latent mean, standard deviation, mean absolute value, and energy,

$$\bar{\mu} = \frac{1}{CK}\sum_{c,k} h_c(t_k), \quad \bar{\sigma}, \quad \overline{|h|} = \frac{1}{CK}\sum_{c,k}|h_c(t_k)|, \quad \bar{E} = \frac{1}{CK}\sum_{c,k} h_c(t_k)^2, \tag{7}$$

yielding $48 + 4 = 52$ descriptors in total. Latent descriptors are compact summaries of the internal INR representation rather than directly interpretable biomechanical variables.

The reconstruction-error family contained 12 descriptors: the global reconstruction R^2, RMSE, and MAE; the RMSE and MAE evaluated within four gait-cycle intervals (0%–20%, 20%–50%, 50%–80%, and 80%–100%); and the absolute error in the total number of extrema, $|N_{\mathrm{ext}}(\hat{x}) - N_{\mathrm{ext}}(x)|$.

For left–right paired descriptors, two bilateral-difference representations were constructed:

$$z_{i,s}^{\mathrm{abs}} = \left|z_{i,s,L} - z_{i,s,R}\right|, \tag{8}$$

and

$$z_{i,s}^{\mathrm{signed}} = z_{i,s,L} - z_{i,s,R}. \tag{9}$$

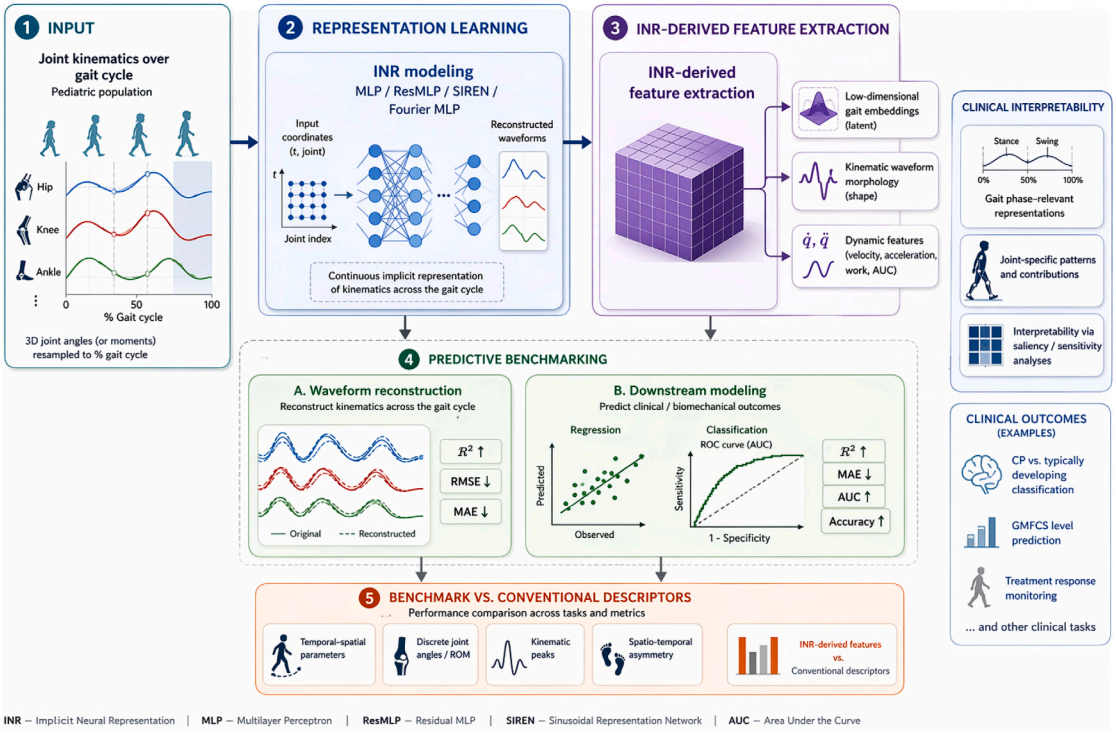

Fig. 1. Overview of the INR-based analytical workflow used in the study. Pediatric gait waveforms were modeled with implicit neural representations, transformed into multiple feature families, and benchmarked in reconstruction, regression, and classification tasks against conventional handcrafted descriptors.

The bilateral-asymmetry representation was defined as their concatenation:

$$z_{i,s}^{\mathrm{asym}} = \left[z_{i,s}^{\mathrm{abs}}, z_{i,s}^{\mathrm{signed}} \right]. \tag{10}$$

The side-specific representation retained the left- and right-side descriptors separately, whereas the bilateral-asymmetry representation retained absolute and signed left–right differences. Both representations therefore contained 1070 features but encoded different predictor values. The complete feature accounting is provided in Table 2.

The main computational workflow is summarized in Fig. 1. A detailed algorithmic description is provided in Appendix A.

The workflow separates three evaluation levels: waveform reconstruction, extraction of INR-derived representations, and downstream developmental prediction.

2.5. Implementation and code structure

All analyses were implemented in Python, with INR models in TensorFlow/Keras and data handling and predictive modeling in NumPy, pandas, SciPy, and scikit-learn. Within each analysis branch, reconstruction metrics and feature definitions were applied consistently; the shallow-MLP and compact Fourier-MLP protocols differed in coordinate encoding, target scaling, architecture, and optimization settings, as detailed above. Full architecture-specific settings and reproducibility information are given in Appendices A.2 and A.3.

2.6. Downstream evaluation protocol

The exploratory architecture-level regression analysis used a fixed 80/20 participant-level train–test split and is reported in Appendix B.3. It is not directly comparable with the primary repeated cross-validation analyses. Continuous chronological age was used as the regression target. For the exploratory classification analyses, chronological age was discretized into three non-overlapping classes: 4–7 years ($4 \le a < 8$), 8–12 years ($8 \le a < 13$), and 13–18 years ($13 \le a \le 18$), where a denotes age in years where (a) denotes age in years. This three-class analytical target was distinct from the four descriptive age categories reported for the full recruited cohort in the Participants subsection. All classification results reported in the main text and appendices correspond to this three-class definition. Data partitioning was performed at the participant level to preserve independence between training and test observations.

The INR feature-ablation analyses and matched representation comparisons used ten repeats of five-fold outer cross-validation. Initial candidate-column construction used target-independent availability and preliminary variance criteria applied once to the complete-case cohort. All subsequent median imputation, winsorization, fold-specific variance filtering, scaling, PCA fitting, model fitting, and hyperparameter selection were performed using the training portion of each cross-validation fold only.

Within each matched representation comparison, identical participants, target definitions, cross-validation partitions, downstream models, and evaluation metrics were used for all representation configurations. Regression performance was quantified using R^2, RMSE, and MAE. Classification performance was assessed primarily using balanced accuracy, macro-F1, and macro-averaged ROC AUC. Weighted F1 was additionally reported for the INR feature-ablation analyses but was not included in the matched representation benchmark.

2.7. INR feature ablations and reconstruction-based architecture mapping

The INR-only ablations used a fixed signal-specific architecture mapping defined before downstream regression and classification. Architecture assignment was based exclusively on waveform-reconstruction quality and did not use chronological-age or age-group information. One common architecture was assigned to both sides of each waveform class.

Fourier-MLP was assigned to ankle dorsiflexion/plantarflexion moment, hip adduction/abduction moment, knee flexion/extension, and hamstring length, whereas SIREN was assigned to pelvic tilt. The

Table 1
Signal-specific INR architectures used in the feature ablations. The architecture for each waveform class was selected from reconstruction performance independently of the downstream age-prediction targets.

Waveform class	Selected INR architecture
Ankle dorsiflexion/plantarflexion moment	Fourier-MLP
Hip adduction/abduction moment	Fourier-MLP
Knee flexion/extension	Fourier-MLP
Hamstring length	Fourier-MLP
Pelvic tilt	SIREN

mapping was subsequently held fixed in all representation-, signal-, and descriptor-family ablations.

For the five selected waveform classes, the side-specific representation contained 1070 features. The bilateral-asymmetry representation contained a corresponding set of 1070 absolute and signed left–right difference descriptors. Their concatenation produced a combined representation of 2140 features.

Three complementary ablation analyses were performed. The representation analysis compared side-specific features, bilateral-asymmetry features, and their concatenation. The signal analysis compared all five selected signals, each individual signal, and leave-one-signal-out configurations. The descriptor-family analysis compared the complete side-specific representation with shape-, latent-, derivative-, energy/AUC-, and reconstruction-error-only subsets.

All ablations used ten repeats of five-fold outer cross-validation. Within each repeat, predictions from the five held-out folds were concatenated to obtain one complete out-of-fold prediction vector for all 73 participants. Regression and classification metrics were then calculated from this complete out-of-fold vector, yielding ten statistically paired repeat-level estimates per feature subset.

For regression, Elastic Net hyperparameters were optimized within an inner cross-validation loop using mean absolute error. The exploratory three-class age-group classification used class-weighted multinomial logistic regression, with the regularization strength and L1/L2 penalty selected by inner stratified cross-validation using balanced accuracy. Missing-value filtering, median imputation, column-wise winsorization at the 1st and 99th training-fold percentiles, variance filtering, and standardization were fitted exclusively on each training fold.

Paired comparisons were performed using Wilcoxon signed-rank tests applied to the ten repeat-level out-of-fold metrics. Holm correction was applied separately within each ablation analysis and evaluation metric. Mean paired differences, bootstrap 95% confidence intervals based on 10,000 resamples, raw and Holm-adjusted p-values, and paired standardized effect sizes (d_z) were calculated.

2.8. Raw-waveform PCA baseline, matched comparisons, and diagnostic analyses

A fold-wise raw-waveform principal component analysis baseline was included as a classical compact waveform representation. Participant-level raw-waveform vectors were formed by concatenating all available signal–side trajectories. Within each training fold, missing values were imputed, predictors were standardized, and PCA was fitted before transformation of the corresponding test fold. The number of components was selected to target 95% cumulative explained variance, subject to a maximum of 30 components. All cross-validation folds reached the predefined cap of 30 PCA components.

The conventional biomechanical feature matrix used in the matched PCA–INR benchmark contained 1695 candidate predictors before fold-specific variance filtering. This matrix differed from the 681-feature biomechanical baseline used in the earlier exploratory architecture-level analysis because the two analyses relied on distinct data-processing and feature-construction pipelines. Accordingly, their feature counts and numerical results should not be interpreted as directly comparable.

The matched representation benchmark compared seven feature configurations: conventional biomechanical descriptors only, raw-waveform PCA only, INR-derived features only, biomechanical descriptors combined with PCA, biomechanical descriptors combined with INR features, PCA combined with INR features, and the full biomechanical+PCA+INR representation.

Chronological-age regression used Elastic Net with internal cross-validation over four L_1/L_2 mixing ratios and 40 logarithmically spaced regularization strengths. Exploratory three-class age-group classification (4–7, 8–12, and 13–18 years) used a Random Forest with 600 trees, class weights recalculated for each bootstrap sample, a minimum leaf size of two, and the square root of the number of predictors considered at each split. The same downstream model specification was applied to all representation configurations within a given task.

The INR-derived feature set in this matched benchmark was generated using the compact Fourier-MLP model described above and followed a feature-construction scheme distinct from the 1070-feature signal-specific representation used in the INR ablation analyses. Descriptors were generated separately for five waveform classes and two sides, yielding ten signal–side feature blocks per participant. Each block initially contained 119 predictors: 32 shape descriptors, eight derivative descriptors, eight energy/AUC descriptors, three reconstruction-quality metrics, 64 latent descriptors, and four training diagnostics. The 64 latent descriptors comprised four summary statistics calculated across the normalized gait-cycle coordinates for each of the 16 units in the final hidden layer. This resulted in 1190 generated predictors. Fourteen globally constant training-diagnostic columns were removed during feature-matrix construction, leaving 1176 candidate predictors before fold-specific missingness and variance filtering. Their composition is reported in Table 3.

The reconstruction diagnostics for this model comprised five selected waveform classes, two sides, and 73 participants, yielding

$$73 \times 5 \times 2 = 730$$

independently fitted waveform models. This diagnostic subset should be distinguished from the complete raw-waveform matrix used for PCA, which contained 128 distinct waveform signals and 256 signal–side pairs. All representation configurations were evaluated using the same 73 participants (Section 2.1), cross-validation partitions, preprocessing rules, targets, and downstream models.

The repeated cross-validation procedure comprised ten repeats of five-fold cross-validation, producing 50 fold-level scores. Descriptive means and standard deviations were calculated from all 50 fold-level results. For statistical inference, however, fold-level scores were first averaged within each repeat, yielding ten paired repeat-level observations. Wilcoxon signed-rank tests were then applied to these ten paired repeat means.

Multiplicity was controlled separately within the regression and classification tasks using the Holm procedure across the twelve planned representation–metric comparisons in each task. Mean paired differences, bootstrap 95% confidence intervals, raw and Holm-adjusted p-values, and paired standardized effect sizes d_z were reported. Fold-level tests were retained only as sensitivity analyses, because folds within repeated cross-validation are not statistically independent. Statistical inference was based on the Holm-adjusted Wilcoxon signed-rank tests. The bootstrap percentile confidence intervals are reported as descriptive summaries of the mean paired difference and were not used for significance decisions. Because these intervals concern the mean difference whereas the Wilcoxon test concerns a rank-based location shift — and because percentile intervals can be anticonservative at the small number of repeat-level observations (n = 10) — the two summaries occasionally diverge; in every such case the corresponding effect was non-significant after Holm correction.

Table 2
Composition of the INR-derived descriptor families under the signal-specific architecture mapping. The side-specific representation concatenated left- and right-side descriptors. The bilateral-asymmetry representation concatenated absolute and signed left–right differences.

Descriptor family	Per signal and side	Five signals, one side	Side-specific L+R	Absolute differences	Signed differences	Asymmetry total
Shape	25	125	250	125	125	250
Latent	52	260	520	260	260	520
Derivative	6	30	60	30	30	60
Energy/AUC	12	60	120	60	60	120
Reconstruction error	12	60	120	60	60	120
Total	107	535	1070	535	535	1070

Table 3
Composition of the compact Fourier-MLP feature set used in the matched PCA–INR benchmark. Descriptors were generated separately for five waveform classes and two sides, yielding ten signal–side feature blocks per participant. The generated-count column reports the feature set before global constant-column removal. Fourteen constant training-diagnostic predictors were removed, leaving 1176 candidate predictors before fold-specific missingness and variance filtering.

Descriptor family	Per signal–side curve	Generated count	Globally constant predictors removed	Retained count
Shape	32	320	0	320
Derivative	8	80	0	80
Energy/AUC	8	80	0	80
Reconstruction quality	3	30	0	30
Latent	64	640	0	640
Training diagnostics	4	40	14	26
Total	**119**	**1190**	**14**	**1176**

Table 4
Chronological-age regression results for alternative bilateral INR feature representations. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations. The INR architecture was selected separately for each waveform class using reconstruction performance.

Representation	Features	R^2	RMSE [years]	MAE [years]
Bilateral asymmetry only	1070	0.245 ± 0.083	3.451 ± 0.185	2.899 ± 0.136
Side-specific only	1070	0.645 ± 0.024	2.369 ± 0.082	1.943 ± 0.061
Side-specific + asymmetry	**2140**	**0.700 ± 0.017**	**2.176 ± 0.064**	**1.801 ± 0.045**

To examine whether reconstruction fidelity translated into developmental predictive relevance, participant-level Fourier-MLP reconstruction quality was correlated with absolute chronological-age prediction error using Spearman correlation. Mean reconstruction R^2, RMSE, and MAE were examined separately. Because the Fourier-MLP reconstruction values were strongly concentrated near their upper limit, the standard deviation, interquartile range, and full range of participant-level reconstruction R^2 were additionally examined to quantify range restriction. Morphology-preservation analyses compared original and reconstructed waveforms using range of motion, peak timing, derivative energy, curvature energy, and extrema-count errors.

3. Results

3.1. Implicit neural representations of gait waveforms

Appendix C.2 and Table C.19 summarize reconstruction performance specifically for the shallow-MLP protocol. Under this protocol, reconstruction quality varied substantially across biomechanical waveform types. The highest mean R^2 values were observed for ankle dorsiflexion/plantarflexion and hip adduction/abduction moment trajectories, whereas pelvic-tilt and hamstring-length trajectories were reconstructed less accurately. These results are specific to the shallow-MLP protocol.

3.2. INR-derived features in chronological age regression

The INR-only regression analysis used the signal-specific architecture mapping (Table 1). The combined side-specific and bilateral-asymmetry representation achieved the strongest performance, with $R^2 = 0.700 \pm 0.017$, RMSE = 2.176 ± 0.064 years, and MAE = 1.801 ± 0.045 years (Table 4).

Relative to the side-specific representation, adding bilateral-asymmetry features increased R^2 by 0.055 and reduced RMSE and MAE by 0.193 and 0.142 years, respectively. All three differences remained significant after Holm correction ($p_{\text{Holm}} = 0.0039$). Bilateral-asymmetry features alone were significantly less predictive than side-specific features ($p_{\text{Holm}} = 0.0039$ for all three regression metrics).

Descriptor-family performance is summarized in Table 5. Energy/AUC descriptors provided the strongest individual-family result, whereas latent and reconstruction-error descriptors showed limited standalone predictive value.

Energy/AUC descriptors significantly outperformed the complete descriptor set: R^2 increased by 0.051, whereas RMSE and MAE decreased by 0.181 and 0.140 years, respectively ($p_{\text{Holm}} = 0.0098$ for all three metrics). Shape-only performance did not differ significantly from the complete descriptor set. Derivative-only performance was nominally lower but did not remain significant after Holm correction.

Leave-one-signal-out and single-signal results are reported in Appendix B.1 (Table B.14); ankle and hip moment features were the strongest incremental contributors.

An exploratory fixed-split comparison of INR architectures is provided in Appendix B.3. These results are not directly comparable with the repeated cross-validation analyses reported above.

3.3. INR-derived feature ablations for age-group classification

The exploratory three-class age-group classification analysis (4–7, 8–12, and 13–18 years) followed the same signal-specific architecture mapping and repeated out-of-fold procedure as the regression analysis. The combined side-specific and bilateral-asymmetry representation achieved the highest mean balanced accuracy and macro-F1 score, whereas the side-specific representation achieved very similar performance with half the number of input features.

Table 5
Chronological-age regression results for INR descriptor families. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations.

Descriptor subset	Features	R^2	RMSE [years]	MAE [years]
All descriptor families	1070	0.645 ± 0.024	2.369 ± 0.082	1.943 ± 0.061
Shape only	250	0.650 ± 0.036	2.350 ± 0.119	1.905 ± 0.110
Latent only	520	−0.077 ± 0.069	4.126 ± 0.134	3.486 ± 0.107
Derivative only	60	0.625 ± 0.032	2.432 ± 0.102	2.009 ± 0.083
Energy/AUC only	**120**	**0.696 ± 0.031**	**2.188 ± 0.110**	**1.803 ± 0.097**
Reconstruction error only	120	0.040 ± 0.034	3.895 ± 0.069	3.324 ± 0.075

Table 6
Exploratory three-class age-group classification results for alternative bilateral INR feature representations. The classes were defined as 4–7, 8–12, and 13–18 years. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations.

Representation	Features	Balanced acc.	Macro-F1	Weighted F1	ROC AUC
Bilateral asymmetry only	1070	0.599 ± 0.042	0.595 ± 0.043	0.604 ± 0.044	0.747 ± 0.035
Side-specific only	1070	0.746 ± 0.029	0.728 ± 0.027	0.721 ± 0.027	0.859 ± 0.021
Side-specific + asymmetry	**2140**	**0.749 ± 0.034**	**0.730 ± 0.038**	**0.726 ± 0.038**	**0.868 ± 0.017**

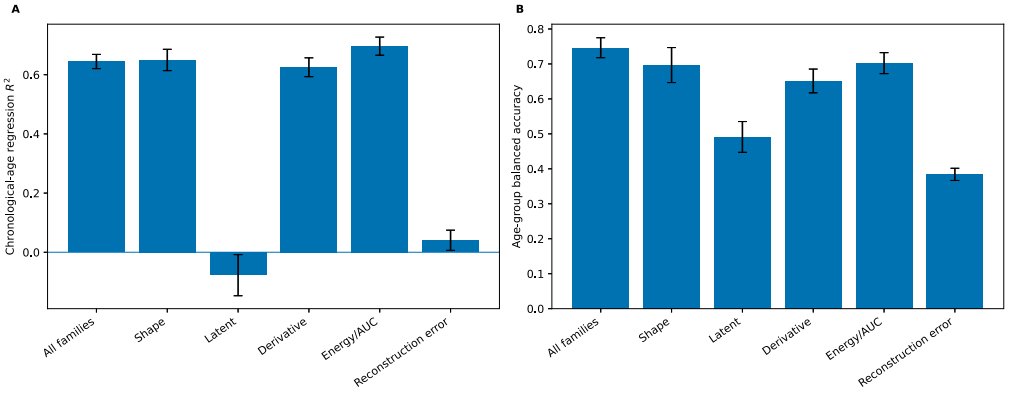

Fig. 2. Performance of INR-derived descriptor families in the two downstream tasks. Bars show mean performance across ten complete repeat-level out-of-fold evaluations. (A) Chronological-age regression quantified using the coefficient of determination (R^2). Energy/AUC descriptors achieved the highest mean regression performance. (B) Exploratory three-class age-group classification (4–7, 8–12, and 13–18 years) quantified using balanced accuracy.

The combined representation yielded only modest mean improvements over the side-specific representation: 0.003 in balanced accuracy, 0.002 in macro-F1, and 0.009 in macro-averaged ROC AUC. Therefore, the two representations should be interpreted as showing broadly comparable classification performance, whereas asymmetry features alone were substantially less informative.

Shape descriptors retained a relatively high ROC AUC, whereas latent and reconstruction-error descriptors were substantially less informative. Complete descriptor-family results are provided in Appendix B.2 (Table B.16).

Signal-level classification results are provided in Appendix B.1 (Table B.15).

Fig. 2 compares the predictive performance of the INR descriptor families. For chronological-age regression, energy/AUC descriptors achieved the highest mean R^2, whereas the complete descriptor representation achieved the highest balanced accuracy for age-group classification.

The INR feature-ablation analyses presented above and the matched PCA–INR representation comparison reported below address complementary questions. The feature-ablation analyses used a signal-specific architecture mapping and evaluated 1070 side-specific features, 1070 bilateral-asymmetry features, and their 2140-feature combination. The matched PCA–INR analysis used a separate compact Fourier-MLP feature representation under common cross-representation evaluation conditions. The former analyses characterize the internal composition of the INR feature space, whereas the latter evaluates the performance of INR-derived features relative to conventional biomechanical descriptors and raw-waveform PCA.

3.4. Matched comparison with raw-waveform PCA and conventional biomechanical descriptors

Table 7 summarizes the matched representation comparison for chronological-age regression. The biomechanics+PCA and biomechanical-only configurations achieved the highest mean performance, with $R^2 = 0.842 \pm 0.066$ and $R^2 = 0.837 \pm 0.070$, respectively. Their difference was small and not statistically significant.

Adding INR-derived features to the biomechanical representation did not significantly change regression performance. The biomechanics+INR configuration achieved $R^2 = 0.833 \pm 0.065$, compared with $R^2 = 0.837 \pm 0.070$ for the biomechanical-only baseline. Standalone INR and raw-waveform PCA representations performed substantially worse than representations containing conventional biomechanical descriptors

Table 8 presents the matched exploratory three-class age-group classification results for the 4–7-, 8–12-, and 13–18-year groups. The biomechanical-only configuration achieved the highest mean balanced accuracy (0.798 ± 0.109), macro-F1 score (0.788 ± 0.114), and ROC AUC (0.910 ± 0.060). Adding raw-waveform PCA features did not significantly change classification performance.

Adding INR-derived features to the biomechanical representation reduced balanced accuracy to 0.767 ± 0.108, macro-F1 to 0.754 ± 0.115, and ROC AUC to 0.896 ± 0.064. Standalone INR-derived features retained measurable age-group information but remained substantially less informative than conventional biomechanical descriptors.

Table 7
Matched comparison of biomechanical, raw-waveform PCA, INR-derived, and combined representations for chronological-age regression. Values are means ± standard deviations across 50 outer-fold evaluations from ten repeats of five-fold cross-validation. Statistical inference was performed using ten paired repeat-level means. Feature counts refer to candidate predictors before fold-specific variance filtering.

Representation	Features	R^2	RMSE [years]	MAE [years]
Biomechanics + PCA	1725	0.842 ± 0.066	1.494 ± 0.281	1.213 ± 0.245
Biomechanics only	1695	0.837 ± 0.070	1.515 ± 0.296	1.223 ± 0.258
Biomechanics + PCA + INR	2901	0.829 ± 0.076	1.552 ± 0.284	1.240 ± 0.236
Biomechanics + INR	2871	0.833 ± 0.065	1.539 ± 0.260	1.230 ± 0.222
INR only	1176	0.147 ± 0.755	3.321 ± 1.318	2.424 ± 0.625
PCA + INR	1206	0.095 ± 0.925	3.351 ± 1.476	2.413 ± 0.673
Raw-waveform PCA only	30	−0.518 ± 2.775	3.587 ± 3.162	2.242 ± 0.979

Table 8
Matched comparison of biomechanical, raw-waveform PCA, INR-derived, and combined representations for exploratory three-class age-group classification. The classes were defined as 4–7, 8–12, and 13–18 years. Values are means ± standard deviations across 50 outer-fold evaluations from ten repeats of five-fold cross-validation. Statistical inference was performed using ten paired repeat-level means. Feature counts refer to candidate predictors before fold-specific variance filtering.

Representation	Features	Balanced accuracy	Macro-F1	ROC AUC
Biomechanics only	1695	0.798 ± 0.109	0.788 ± 0.114	0.910 ± 0.060
Biomechanics + PCA	1725	0.784 ± 0.109	0.772 ± 0.117	0.905 ± 0.061
Biomechanics + PCA + INR	2901	0.768 ± 0.116	0.756 ± 0.123	0.898 ± 0.064
Biomechanics + INR	2871	0.767 ± 0.108	0.754 ± 0.115	0.896 ± 0.064
INR only	1176	0.648 ± 0.116	0.625 ± 0.118	0.836 ± 0.059
PCA + INR	1206	0.642 ± 0.105	0.615 ± 0.109	0.837 ± 0.062
Raw-waveform PCA only	30	0.552 ± 0.111	0.524 ± 0.127	0.781 ± 0.081

3.5. Formal paired comparisons between representations

Paired comparisons based on the ten cross-validation repeat means confirmed that adding PCA to the biomechanical representation did not significantly alter regression or classification performance after Holm correction.

For chronological-age regression, adding INR-derived features to the biomechanical baseline changed mean R^2 by −0.0039, RMSE by 0.0240 years, and MAE by 0.0078 years. None of these differences was statistically significant after Holm correction ($p_{\text{Holm}} = 1.000$ for all three metrics).

For exploratory three-class age-group classification, adding INR-derived features significantly reduced balanced accuracy by 0.0304 ($p_{\text{Holm}} = 0.0307$), macro-F1 by 0.0347 ($p_{\text{Holm}} = 0.0234$), and ROC AUC by 0.0134 ($p_{\text{Holm}} = 0.0234$). Standalone INR and standalone raw-waveform PCA representations were significantly worse than the biomechanical baseline for all evaluated regression and classification metrics ($p_{\text{Holm}} = 0.0234$).

Complete paired comparisons, including bootstrap confidence intervals, effect sizes, raw p-values, and Holm-adjusted p-values, are provided in Appendix C.3, Tables C.20 and C.21.

3.6. Auxiliary reconstruction diagnostics

Participant-level reconstruction quality was not significantly associated with absolute chronological-age prediction error. Because reconstruction R^2 values were strongly concentrated near one, the analysis was affected by severe range restriction and should be interpreted descriptively. Complete correlation results and the corresponding visualization are provided in Appendix C.4.

The compact Fourier-MLP protocol achieved near-perfect in-sample approximation. Global amplitude-related properties were generally preserved, whereas derivative- and extrema-based waveform characteristics were more sensitive to smoothing. Complete morphology-preservation and convergence diagnostics are provided in Appendix C.5.

3.7. Benchmark summary

The representation-, signal-, and descriptor-family ablations are treated as secondary INR-internal analyses, whereas the matched PCA–INR benchmark remains the primary comparison between representation families.

Despite this high approximation fidelity, Fourier-MLP-derived features showed limited standalone utility for chronological age regression and did not provide an incremental advantage over conventional biomechanical descriptors. Selected INR-derived feature families retained some information relevant to exploratory age-group classification, but the biomechanical baseline remained stronger, and adding INR features significantly reduced classification performance in the matched repeat-level comparisons.

Table 9 summarizes the principal representation analyses.

4. Discussion

The present study evaluated implicit neural representations as continuous encodings of pediatric gait waveforms for reconstruction and developmental modeling. The INR-only ablations showed that reconstruction-selected INR features retained substantial age-related information: combining side-specific and bilateral-asymmetry descriptors yielded $R^2 = 0.700 \pm 0.017$ for chronological-age regression. Nevertheless, representations containing conventional biomechanical descriptors achieved the strongest performance in the matched cross-representation benchmark, and the biomechanical-only baseline was statistically indistinguishable from the highest-performing biomechanics+PCA configuration.

A first key observation was that reconstruction fidelity depended on both waveform type and fitting protocol. Under the shallow-MLP protocol, ankle and hip moment trajectories were reconstructed more accurately than pelvic-tilt and hamstring-length trajectories. In contrast, the compact Fourier-MLP protocol achieved near-perfect in-sample approximation across the same selected waveform classes.

This contrast indicates that the lower reconstruction values in Table C.19 should not be interpreted as an intrinsic limitation of INR modeling for these signals. Rather, they reflect the capacity, coordinate representation, scaling, and optimization settings of the shallow-MLP

Table 9
Overview of the primary waveform-representation analyses. Each row reports a single headline metric (R^2 for regression, balanced accuracy for classification, and explained variance for the PCA encoding). Complete metrics, including standard deviations, are provided in the source tables listed in the penultimate column. The INR feature analyses characterize the internal predictive utility of the signal-specific INR feature space, whereas the matched comparisons evaluate INR- and PCA-derived features relative to conventional biomechanical descriptors.

Analysis	Representation	Headline result	Source	Interpretation
Raw-waveform encoding	30-component fold-wise PCA	Explained variance = 0.899	Table C.18	Most waveform variance was preserved, but variance preservation did not guarantee superior downstream prediction.
Regression, INR ablation	Side-specific + asymmetry	$R^2 = 0.700$	Table 4	Strongest INR-only feature configuration.
Classification, INR ablation	Side-specific + asymmetry	Balanced accuracy = 0.749	Table 6	Highest mean INR-only classification result.
Regression, matched	Biomechanics + PCA	$R^2 = 0.842$	Table 7	Highest mean regression result, although the difference from biomechanics alone was not statistically significant.
Regression, matched	Biomechanics only	$R^2 = 0.837$	Table 7	Strong and stable conventional baseline.
Regression, matched	Biomechanics + INR	$R^2 = 0.833$	Table 7	Adding INR-derived features did not significantly change performance.
Regression, matched	INR only	$R^2 = 0.147$	Table 7	Weak and unstable standalone performance.
Classification (3 classes), matched	Biomechanics only	Balanced accuracy = 0.798	Table 8	Highest mean classification performance.
Classification (3 classes), matched	Biomechanics + PCA	Balanced accuracy = 0.784	Table 8	Performance did not differ significantly from that of biomechanics alone.
Classification (3 classes), matched	Biomechanics + INR	Balanced accuracy = 0.767	Table 8	Adding INR-derived features significantly reduced the evaluated classification metrics.
Classification (3 classes), matched	INR only	Balanced accuracy = 0.648	Table 8	Retained measurable age-group information but remained weaker than the biomechanical representation.
Reconstruction–prediction relation	Fourier-MLP reconstruction versus prediction error	Weak, non-significant association	Table C.22	Near-perfect in-sample reconstruction did not ensure accurate prediction; the analysis was auxiliary and descriptive.

protocol. At the same time, the near-perfect Fourier-MLP results demonstrate approximation capacity at the sampled coordinates and do not by themselves establish that the resulting representation preserves developmentally informative variation.

A second important finding was that signal-architecture compatibility was important for constructing informative INR-derived feature spaces. Fourier-MLP was assigned to four waveform classes and SIREN to pelvic tilt (Table 1). The resulting signal-specific mapping retained measurable downstream predictive information, indicating that a single architecture may not be equally suitable for kinetic, kinematic, and muscle-length trajectories with different waveform characteristics [13,14].

Despite the informative INR-only results, conventional biomechanical descriptors remained more accurate for chronological-age prediction, and adding INR-derived features to the biomechanical representation did not provide a consistent incremental advantage. INR-derived descriptors can therefore encode developmental information without necessarily outperforming physically interpretable biomechanical features, which is consistent with previous work showing that carefully designed biomechanical features remain competitive with data-driven approaches in gait analysis [2,24,25].

The added raw-waveform PCA baseline clarifies the relationship between INR-derived representations and established compact waveform encodings. Although the 30-component PCA representation preserved approximately 90% of the raw waveform variance, PCA-only features did not outperform conventional biomechanical descriptors in chronological age regression or exploratory age-group classification. This indicates that preservation of dominant waveform variance is not equivalent to preservation of developmentally predictive information. Similarly, the Fourier-MLP model approximated waveforms almost perfectly but did not provide consistent incremental value beyond conventional biomechanical descriptors.

Conceptually, PCA and functional PCA provide cohort-level linear decompositions of waveform variability, whereas the INR models used here provide participant- and signal-specific continuous coordinate-to-amplitude mappings [9,10,26]. Thus, PCA emphasizes dominant modes of variance shared across participants, while INRs emphasize smooth functional approximation of individual trajectories. Functional data analysis and fPCA may therefore be more directly aligned with population-level developmental modeling, whereas INRs provide advantages related to continuous interpolation, differentiability, compact signal storage, and morphology-preserving reconstruction. The present results show that these properties do not automatically translate into stronger age prediction.

The morphology-preservation analysis also helps explain the limited downstream utility of INR-derived features. INR reconstructions preserved global amplitude-related properties relatively well, but local timing, derivative, curvature, and extrema-based characteristics were more sensitive to smoothing. In pediatric gait analysis, such local waveform features may contain participant-specific or developmentally relevant information that is not fully captured by reconstruction objectives alone. Therefore, future INR-based approaches may require task-aware, shared, or meta-learned representations rather than independently fitted participant-level models optimized only for waveform reconstruction.

Because the present implementation fitted participant-specific INR models independently, the resulting parameter spaces are not guaranteed to be aligned across participants. Future work should therefore evaluate shared, amortized, or meta-learned INR frameworks, including functa-style representations in which fitted functions are treated as data objects for downstream representation learning [30]. Such approaches may improve feature comparability, computational efficiency, and adaptation to population-level prediction tasks. Partitioned or piecewise INR strategies may provide an additional route to improving fitting efficiency for signals with heterogeneous local structure by assigning smaller implicit models to separate signal regions [31]. In gait analysis, a related strategy could involve phase-specific INR components corresponding to stance, pre-swing, and swing; however, the utility of such partitioning for biomechanical waveforms remains to be evaluated.

The descriptor-family analysis showed that age-related information was not distributed uniformly across INR-derived feature types, and that the most informative subset depended on the downstream

task. For continuous chronological-age regression, the compact energy/AUC subset provided the strongest representation and significantly outperformed the complete descriptor set ($R^2 = 0.696 \pm 0.031$ versus 0.645 ± 0.024; $p_{\mathrm{Holm}} = 0.0098$), whereas shape-only performance was statistically indistinguishable from the complete set and derivative-only performance was nominally lower but not significant after Holm correction. For exploratory age-group classification the ordering differed: the complete descriptor set achieved the highest balanced accuracy (0.746 ± 0.029), the energy/AUC subset did not retain its advantage (0.702 ± 0.030), and shape descriptors retained a comparatively high ROC AUC (0.847 ± 0.030) despite using only 250 features. Latent and reconstruction-error descriptors were weakly predictive in both tasks. Integral, physically interpretable summaries of the fitted waveform therefore carried more developmental information than summaries of the internal network state, although the most informative INR summary depended on the downstream task and evaluation metric. From a translational perspective, INRs are best viewed as continuous representations of waveform morphology rather than direct replacements for conventional biomechanical descriptors. Their potential advantages include smooth interpolation, access to waveform derivatives, potentially compact parametric signal representation, flexible morphology-based feature construction, and initialization of future task-aware representation-learning or phenotype-discovery pipelines [13–15].

5. Limitations of the study

The present findings should be interpreted cautiously. The cohort size was modest relative to the dimensionality of several INR-derived feature sets, which may have limited model stability and generalizability despite the use of regularization and cross-validation, particularly for the exploratory three-class age-group classification.

Chronological age was used only as a practical proxy for development rather than a complete measure of neuromuscular maturation. All repeated cross-validation analyses should therefore be interpreted as complete-case analyses on 73 of the 78 recruited participants. The constrained PCA representation retained 0.899 ± 0.017 of the training-fold variance rather than the targeted 95%.

An additional limitation concerns the interpretation of reconstruction quality. The shallow-MLP and Fourier-MLP results were obtained under different architectures, target-scaling procedures, and optimization settings and therefore should not be treated as directly interchangeable estimates. Moreover, reconstruction metrics were calculated at the same 101 gait-cycle coordinates used to fit each participant-specific INR. The reported reconstruction values consequently quantify in-sample approximation rather than interpolation or generalization to held-out coordinates. The reconstruction–prediction correlation analysis was also affected by severe range restriction because participant-level Fourier-MLP reconstruction R^2 values were concentrated close to one. Therefore, the absence of a significant correlation should not be interpreted as definitive evidence that reconstruction fidelity and developmental utility are statistically independent.

The signal-specific architecture mapping was defined on the complete-case cohort and not repeated within each outer training fold; it should therefore be interpreted as a fixed representation-design decision rather than fully nested model selection. In addition, the feature-ablation analyses were restricted to five selected waveform classes. Validation using independently prespecified waveform sets and larger external cohorts is required.

The large fold-level variability of the standalone raw-waveform PCA and INR-based regression models was accompanied by occasional extrapolations outside the observed pediatric age range. This instability further supports the interpretation that these standalone representations were insufficiently regularized for reliable standalone age prediction in the present small-sample cohort. A further limitation concerns the initial construction of the candidate predictor set. Target-independent availability and preliminary variance criteria were applied once to the complete-case cohort before repeated cross-validation. Although these criteria did not use chronological-age or age-group labels, they incorporated information about the predictor distribution across the full analytical cohort, including observations subsequently assigned to test folds. This procedure may have introduced a small optimistic bias relative to a fully fold-wise feature-construction pipeline. Future analyses should repeat all availability- and variance-based screening exclusively within each outer training fold.

6. Conclusions

Implicit neural representations provided continuous coordinate-based encodings of selected pediatric gait waveforms, although the resulting downstream descriptor matrices remained high-dimensional. Reconstruction fidelity depended strongly on the fitting protocol. The shallow-MLP protocol showed substantial signal-dependent variation in reconstruction quality, whereas the compact Fourier-MLP protocol achieved near-perfect in-sample approximation of the same selected waveform classes. Despite the high approximation fidelity of the Fourier-MLP model, the matched representation benchmark showed that neither raw-waveform PCA nor INR-derived representations replaced conventional biomechanical descriptors for developmental prediction in this cohort. PCA preserved most raw waveform variance, but PCA-derived features did not provide a robust incremental improvement over biomechanical descriptors. These findings reinforce the conclusion that waveform approximation fidelity, variance preservation, and developmental predictive utility are distinct representation properties.

In the INR feature analyses, signal-specific architecture selection and the combination of side-specific and bilateral-asymmetry descriptors retained substantial developmental information. The combined representation achieved $R^2 = 0.700 \pm 0.017$ for chronological-age regression and a balanced accuracy of 0.749 ± 0.034 for exploratory age-group classification. Energy/AUC descriptors provided the strongest individual regression feature family, whereas the complete descriptor set achieved the highest overall classification performance.

Conventional biomechanical descriptors nevertheless remained more accurate in the matched representation comparison, and adding INR-derived features did not provide a consistent incremental advantage. INRs should therefore be viewed as complementary continuous waveform representations whose utility depends on signal-architecture compatibility and feature design, rather than as direct replacements for established biomechanical descriptors. Their practical value may therefore lie in applications such as waveform compression, interpolation of sparsely sampled gait cycles, derivative-based morphology analysis, and initialization of task-aware representation-learning pipelines, rather than in incremental gains in developmental prediction.

CRediT authorship contribution statement

Agnieszka Pregowska: Writing – original draft, Validation, Software, Methodology. **Konrad Pauk:** Writing – review & editing, Visualization, Resources, Formal analysis. **Jacek Tutak:** Writing – review & editing, Visualization, Methodology, Formal analysis. **Jolanta Pauk:** Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization.

Code availability statement

Custom code used for waveform preprocessing, INR modeling, feature extraction, and downstream predictive analyses is available from the corresponding author upon reasonable request, subject to institutional and project-related constraints.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank James Raso and the clinical staff for their support and assistance.

Appendix A. Implementation and reproducibility details

A.1. Computational pipeline

For clarity and reproducibility, the main computational workflow is summarized in Algorithm 1. The pipeline includes preprocessing, INR fitting, feature extraction, and downstream regression and classification analyses used to benchmark INR-derived representations against conventional biomechanical descriptors.

Algorithm 1 INR-based gait waveform analysis pipeline

Require: Selected gait waveforms for all subjects
Ensure: Reconstruction metrics, INR-derived feature sets, downstream regression/classification results
for each waveform type and side **do**
  Normalize waveform coordinates over the gait cycle
  Train INR model to map coordinate → waveform value
  Compute reconstruction metrics (R^2, RMSE, MAE)
  Extract INR-derived feature families:
    shape descriptors
    latent descriptors
    derivative-based descriptors
    energy/AUC descriptors
    reconstruction-error descriptors
  Construct side-specific and bilateral-asymmetry representations
end for
Branch A (INR feature ablations): assign one architecture to each waveform class using reconstruction quality
  Construct side-specific, bilateral-asymmetry, signal, and descriptor-family subsets
  Evaluate INR-only feature sets in chronological age regression and age-group classification
Branch B (matched PCA–INR benchmark): construct a separate compact Fourier-MLP representation
  Evaluate biomechanical, raw-waveform PCA, INR-derived, and combined representations under common cross-validation
Compare INR-derived features with conventional handcrafted descriptors and raw-waveform PCA
Summarize reconstruction, regression, and classification benchmarks

A.2. INR architectures and training settings

See Tables A.10–A.13.

A.3. Reproducibility and code structure

All analyses were implemented in Python within a unified computational pipeline integrating waveform preprocessing, INR fitting, feature extraction, and downstream predictive modeling. INR models were implemented in TensorFlow/Keras, whereas data handling and machine-learning analyses were performed using NumPy, pandas, SciPy, and scikit-learn.

Initial target-independent candidate-column construction used availability and variance criteria applied to the complete-case cohort. Fold-specific median imputation, winsorization where applicable, scaling, PCA, model fitting, and hyperparameter selection where applicable were performed using training data only. Randomness associated with model initialization and data partitioning was controlled using fixed random seeds.

The compared INR architectures and shared implementation settings are summarized in Tables A.10 and A.11. Representative hyperparameters and optimization settings are provided in Tables A.12 and A.13.

The compact Fourier-MLP model used in the matched PCA–INR benchmark was a separate diagnostic implementation from the higher-capacity Fourier MLP listed in Tables A.12 and A.13. The diagnostic implementation used 48–16 ReLU-activated hidden layers, per-curve target standardization, a maximum of 350 epochs, and an early-stopping patience of 30 epochs. Code availability is specified in the Code Availability Statement.

Appendix B. Supplementary INR feature analyses

B.1. Signal-specific regression and classification analyses

The leave-one-signal-out analysis identified ankle and hip moments as the strongest incremental contributors. Removing either signal significantly reduced R^2 and increased RMSE and MAE ($p_{\text{Holm}} = 0.0195$ for all three metrics). Removing hamstring-length features significantly reduced R^2 and increased RMSE, whereas the corresponding MAE difference was not significant after Holm correction. Removal of knee flexion/extension or pelvic-tilt features did not result in a statistically detectable change relative to the complete five-signal representation. Nevertheless, every single-signal model performed significantly worse than the complete five-signal representation.

The architecture was assigned separately to each waveform class using reconstruction performance; therefore, a separate downstream architecture ranking was not performed for this analysis.

B.2. Descriptor-family analysis for three-class age-group classification

The complete descriptor representation achieved the highest balanced accuracy and macro-F1 score. Shape descriptors retained a relatively high ROC AUC despite using only 250 features, whereas latent and reconstruction-error descriptors were substantially less informative. This ranking differed from the regression analysis, in which the compact energy/AUC subset provided the strongest individual-family performance.

B.3. Exploratory architecture-level regression analysis

The 681-feature biomechanical baseline reported in this exploratory analysis was generated using the feature matrix available to the architecture–comparison pipeline. It was not identical to the expanded 1695-candidate-predictor biomechanical representation used in the subsequent matched PCA–INR benchmark. Accordingly, the exploratory analysis is reported separately and is not used for direct inference about the matched representation comparison.

As shown in Table B.17, the effect of adding INR-derived features to biomechanical descriptors varied across architectures. The highest performance in a single train–test split was observed for `ElasticNet_BiomechPlus_ResMLP` ($R^2 = 0.841$). However, this improvement was not consistently reflected in cross-validation results, where the highest mean performance was achieved by `ElasticNet_BiomechPlus_MLP_Deep` (CV mean $R^2 = 0.786$).

Importantly, this performance was comparable to the biomechanical baseline without INR features (CV mean $R^2 = 0.785$), indicating that

Table A.10
Overview of the implicit neural representation (INR) architectures and shared training configuration used in the study.

Architecture	Input representation	Core mechanism	Output	Main purpose in study
MLP (Shallow)	Normalized gait-cycle coordinate	Fully connected feedforward network	Continuous waveform value	Baseline implicit waveform approximation
MLP (Deep)	Normalized gait-cycle coordinate	Deeper fully connected network	Continuous waveform value	Higher-capacity implicit representation
ResMLP	Normalized gait-cycle coordinate	Residual fully connected blocks	Continuous waveform value	Improved optimization stability and representational depth
SIREN	Normalized gait-cycle coordinate	Sinusoidal activations	Continuous waveform value	Frequency-sensitive implicit representation
Fourier MLP	Fourier-encoded coordinate input	Frequency-enhanced coordinate embedding + MLP	Continuous waveform value	Improved representation of oscillatory waveform structure

Table A.11
Shared training and evaluation settings for INR waveform modeling.

Parameter	Setting
Input domain	Time-normalized gait-cycle coordinate (0 to 1)
Prediction target	Scalar waveform value at each normalized coordinate
Waveform granularity	100% gait cycle representation
Training objective	Waveform reconstruction
Primary reconstruction metrics	R^2, RMSE, MAE
Architecture families compared	MLP, Deep MLP, ResMLP, SIREN, Fourier MLP
Feature extraction outputs	Shape, latent, derivative, energy/AUC, reconstruction-error, side-specific, bilateral-asymmetry, and signal-specific feature subsets
Downstream tasks	Chronological-age regression and exploratory three-class age-group classification (4–7, 8–12, and 13–18 years)
Comparative baseline	Conventional handcrafted biomechanical waveform descriptors
Evaluation design	Held-out testing and/or repeated cross-validation, depending on task

Table A.12
Representative implementation parameters for the architectures used in the exploratory architecture benchmark. The Fourier MLP row refers to the higher-capacity exploratory variant with three 64-unit hidden layers and should not be confused with the compact Fourier-MLP protocol used in the matched PCA–INR benchmark. The compact protocol is described separately in Appendix A.3.

Architecture	Hidden layers	Hidden units	Latent dim	Activation	Coordinate encoding	Notes
MLP (Shallow)	3	32–8–32	8	tanh	None	Shallow-MLP protocol
MLP (Deep)	4	64	12	tanh	None	Increased depth
ResMLP	4	64	12	tanh	None	Residual connections
SIREN	3	64	12	sine	None	Periodic activations, $\omega_0 = 30.0$
Fourier MLP (higher-capacity)	3	64	12	ReLU	Fourier features	Fourier dim = 16, $\sigma = 8.0$

Table A.13
Architecture-specific optimization settings used in the exploratory INR architecture benchmark. The Fourier MLP settings in this table refer to the higher-capacity exploratory variant. The compact Fourier-MLP protocol used in the matched PCA–INR benchmark followed a separate training configuration described in Appendix A.3.

Architecture	Epochs	Batch size	Patience	Learning rate
MLP (Shallow)	500	32	40	1×10^{-3}
MLP Deep	500	32	40	1×10^{-3}
Fourier MLP (higher-capacity)	500	32	40	1×10^{-3}
SIREN	600	32	50	5×10^{-4}
ResMLP	500	32	40	1×10^{-3}

the inclusion of INR-derived features did not result in a consistent improvement in generalization performance. The discrepancy between single-split and cross-validated results suggests that isolated performance gains may be sensitive to data partitioning and should therefore be interpreted cautiously.

Appendix C. Matched PCA–INR benchmark and reconstruction diagnostics

The matched waveform-level benchmark included 73 participants with conventional biomechanical descriptors, raw-waveform PCA features, and INR-derived features. The source waveform inventory comprised 128 waveform labels and 256 possible signal–side combinations. Across the 73 participants, 18,542 participant-level signal–side trajectories were available, each represented by 101 normalized gait-cycle samples. This yielded

$$18,542 \times 101 = 1,872,742$$

observed waveform values. The nominal inventory comprised

$$73 \times 256 = 18,688$$

possible participant-signal–side trajectories. Therefore, 146 complete trajectories were unavailable in the source data. The fold-wise PCA representation contained 30 components and retained 0.899 ± 0.017 of the waveform variance, with fold-level values ranging from 0.861 to 0.913. All folds reached the predefined 30-component cap; thus, the

Table B.14
Chronological-age regression results for signal-specific and leave-one-signal-out INR feature subsets. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations.

Signal subset	Features	R^2	RMSE [years]	MAE [years]
All five selected signals	1070	0.645 ± 0.024	2.369 ± 0.082	1.943 ± 0.061
All except ankle moment	856	0.535 ± 0.029	2.711 ± 0.084	2.317 ± 0.086
All except hip moment	856	0.557 ± 0.038	2.644 ± 0.114	2.179 ± 0.098
All except knee flexion/extension	856	0.653 ± 0.030	2.339 ± 0.100	1.913 ± 0.082
All except hamstring length	856	0.616 ± 0.027	2.462 ± 0.089	1.982 ± 0.070
All except pelvic tilt	856	0.635 ± 0.026	2.402 ± 0.084	1.934 ± 0.088
Ankle moment only	214	0.514 ± 0.066	2.768 ± 0.181	2.210 ± 0.130
Hip moment only	214	0.411 ± 0.057	3.050 ± 0.149	2.502 ± 0.126
Knee flexion/extension only	214	0.247 ± 0.041	3.449 ± 0.093	2.886 ± 0.095
Hamstring length only	214	0.216 ± 0.047	3.519 ± 0.104	3.041 ± 0.087
Pelvic tilt only	214	0.038 ± 0.088	3.897 ± 0.172	3.353 ± 0.110

Table B.15
Exploratory three-class age-group classification performance of signal-specific and leave-one-signal-out INR feature subsets. The classes were defined as 4–7, 8–12, and 13–18 years. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations.

Signal subset	Features	Balanced acc.	Macro-F1	Weighted F1	ROC AUC
All five selected signals	1070	0.746 ± 0.029	0.728 ± 0.027	0.721 ± 0.027	0.859 ± 0.021
All except ankle moment	856	0.682 ± 0.026	0.669 ± 0.025	0.661 ± 0.026	0.831 ± 0.026
All except hip moment	856	0.712 ± 0.024	0.694 ± 0.030	0.691 ± 0.030	0.850 ± 0.019
All except knee flexion/extension	856	0.705 ± 0.021	0.675 ± 0.029	0.664 ± 0.028	0.845 ± 0.013
All except hamstring length	856	0.724 ± 0.028	0.704 ± 0.028	0.696 ± 0.029	0.850 ± 0.018
All except pelvic tilt	856	0.736 ± 0.039	0.713 ± 0.043	0.709 ± 0.043	0.865 ± 0.024
Ankle moment only	214	0.666 ± 0.043	0.642 ± 0.038	0.640 ± 0.038	0.820 ± 0.038
Hip moment only	214	0.526 ± 0.037	0.491 ± 0.041	0.481 ± 0.042	0.693 ± 0.034
Knee flexion/extension only	214	0.573 ± 0.037	0.561 ± 0.039	0.563 ± 0.040	0.724 ± 0.021
Hamstring length only	214	0.533 ± 0.035	0.509 ± 0.037	0.501 ± 0.036	0.683 ± 0.027
Pelvic tilt only	214	0.484 ± 0.043	0.469 ± 0.044	0.466 ± 0.045	0.636 ± 0.035

Table B.16
Exploratory three-class age-group classification results for INR descriptor families. The classes were defined as 4–7, 8–12, and 13–18 years. Values are means ± standard deviations across ten complete repeat-level out-of-fold evaluations.

Descriptor subset	Features	Balanced acc.	Macro-F1	Weighted F1	ROC AUC
All descriptor families	1070	0.746 ± 0.029	0.728 ± 0.027	0.721 ± 0.027	0.859 ± 0.021
Shape only	250	0.697 ± 0.050	0.680 ± 0.053	0.672 ± 0.055	0.847 ± 0.030
Latent only	520	0.491 ± 0.044	0.492 ± 0.043	0.495 ± 0.043	0.655 ± 0.041
Derivative only	60	0.651 ± 0.034	0.634 ± 0.037	0.620 ± 0.038	0.776 ± 0.026
Energy/AUC only	120	0.702 ± 0.030	0.673 ± 0.028	0.666 ± 0.029	0.807 ± 0.028
Reconstruction error only	120	0.384 ± 0.017	0.384 ± 0.018	0.383 ± 0.025	0.550 ± 0.022

Table B.17
Elastic Net age regression with biomechanical descriptors and selected INR architectures. The Fourier MLP in this table refers to the higher-capacity exploratory architecture (three hidden layers, 64 units; Table A.12), not the compact Fourier-MLP protocol used in the matched PCA–INR benchmark.

Model	Features	R^2	RMSE	MAE	CV mean R^2
Biomechanics + MLP (Deep)	1751	0.792	1.839	1.464	0.786
Biomechanics only	681	0.809	1.761	1.449	0.785
Biomechanics + ResMLP	1751	0.841	1.607	1.375	0.776
Biomechanics + SIREN	1751	0.825	1.684	1.410	0.762
Biomechanics + Fourier MLP	1751	0.821	1.704	1.358	0.745
Biomechanics + MLP (Shallow)	1591	0.802	1.791	1.458	0.628

constrained PCA representation preserved most, but not 95%, of the raw-waveform variance.

C.1. Matched waveform-level dataset

See Table C.18.

C.2. Original shallow-MLP reconstruction and Fourier-MLP diagnostic examples

Table C.19 reports shallow-MLP reconstruction performance.

Representative original waveforms and their reconstructions obtained using the compact Fourier-MLP protocol are shown in Fig. C.3. The displayed pelvic-tilt and knee-flexion trajectories were approximated with R^2 values close to one. These examples demonstrate the high in-sample approximation capacity of the Fourier-encoded model at the coordinates used for fitting.

C.3. Paired statistical comparisons

See Tables C.20 and C.21.

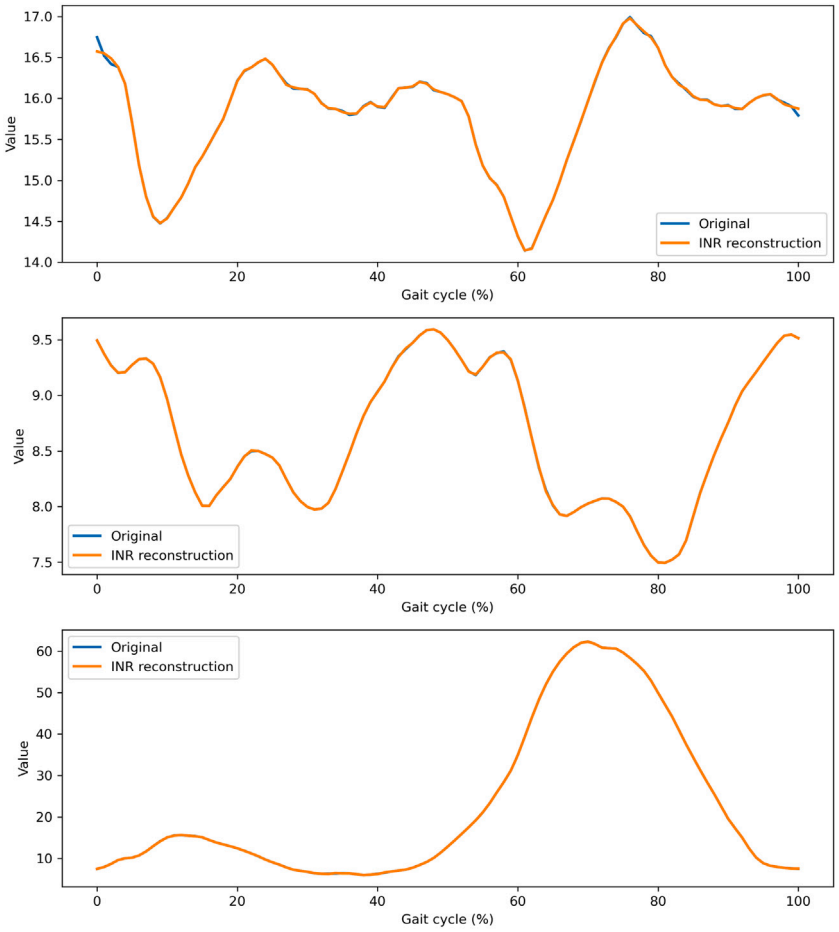

Fig. C.3. Representative examples of original gait waveforms and their reconstructions obtained using the compact Fourier-MLP protocol. The pelvic-tilt and knee-flexion trajectories were approximated with R^2 values close to one at the 101 gait-cycle coordinates used for model fitting. The figure therefore illustrates high in-sample approximation fidelity rather than performance at independently held-out coordinates.

Table C.18
Matched waveform-level data used in the PCA–INR representation analysis.

Quantity	Value
Participants in the full cohort	78
Participants in the matched representation benchmark	73
Distinct waveform labels	128
Possible signal–side combinations	256
Possible participant–signal–side trajectories	18,688
Observed participant–signal–side trajectories	18,542
Unavailable participant–signal–side trajectories	146
Observed waveform values	1,872,742
PCA components retained per fold	30
Mean explained waveform variance	0.899 ± 0.017
Range of explained waveform variance	0.861–0.913

Table C.19
Reconstruction performance of the original shallow-MLP INR baseline for the five selected gait waveforms. The model used three tanh-activated hidden layers of sizes 32–8–32 and was fitted directly to waveform amplitudes without per–curve target standardization. Values are means across 73 participants.

Signal	Side	Mean R^2	Mean RMSE	Mean MAE
Ankle dorsiflexion/plantarflexion moment	Left	0.815	0.138	0.112
Hip adduction–abduction moment	Left	0.745	0.112	0.089
Ankle dorsiflexion/plantarflexion moment	Right	0.743	0.162	0.132
Hip adduction–abduction moment	Right	0.739	0.112	0.089
Knee flexion–extension	Left	0.532	11.643	8.734
Knee flexion–extension	Right	0.508	12.101	9.017
Hamstring length	Left	0.368	0.020	0.016
Hamstring length	Right	0.365	0.020	0.016
Pelvic tilt	Right	0.100	0.760	0.640
Pelvic tilt	Left	0.074	0.772	0.648

C.4. Relationship between INR reconstruction fidelity and developmental prediction

Participant-level reconstruction quality from the compact Fourier-MLP model was correlated with absolute chronological age-prediction error. All associations were weak and statistically non-significant. For example, mean reconstruction R^2 was not significantly associated with prediction error for either the biomechanical+INR model ($\rho = -0.050$, $p = 0.676$) or the INR-only model ($\rho = -0.142$, $p = 0.230$).

However, the Fourier-MLP reconstruction R^2 values exhibited severe range restriction. Across participants, mean participant-level reconstruction R^2 was 0.999978, with a standard deviation of 0.000015, an interquartile range of 0.000008, and a total range of 0.000117. Consequently, the absence of a significant correlation should be interpreted as a descriptive auxiliary result rather than strong evidence of statistical independence between reconstruction fidelity and developmental relevance (see Fig. C.4 and Table C.23).

Table C.20
Selected repeat-level paired statistical comparisons for chronological age regression. Differences are calculated as comparator minus biomechanical baseline. Wilcoxon signed-rank tests were applied to ten paired cross-validation repeat means. Holm adjustment was performed across twelve planned comparisons within the regression task.

Comparator	Metric	Mean difference	95% CI	p_{raw}	p_{Holm}	d_z
Biomechanics + PCA	R^2	+0.0051	[0.0001, 0.0131]	0.1309	0.7852	0.440
Biomechanics + PCA	RMSE	−0.0208	[−0.0550, 0.0016]	0.1602	0.8008	−0.415
Biomechanics + PCA	MAE	−0.0099	[−0.0329, 0.0071]	0.6250	1.0000	−0.284
Biomechanics + INR	R^2	−0.0039	[−0.0138, 0.0077]	0.4316	1.0000	−0.209
Biomechanics + INR	RMSE	+0.0240	[−0.0267, 0.0688]	0.3750	1.0000	0.291
Biomechanics + INR	MAE	+0.0078	[−0.0515, 0.0569]	0.6953	1.0000	0.084
INR only	R^2	−0.6899	[−0.8161, −0.5699]	0.0020	0.0234	−3.235
Raw-waveform PCA only	R^2	−1.3545	[−1.7542, −1.0001]	0.0020	0.0234	−2.142

Table C.21
Selected repeat-level paired statistical comparisons for exploratory three-class age-group classification (4–7, 8–12, and 13–18 years). Differences are calculated as comparator minus the biomechanical baseline. Wilcoxon signed-rank tests were applied to ten paired cross-validation repeat means. Holm adjustment was performed across twelve planned comparisons within the classification task.

Comparator	Metric	Mean difference	95% CI	p_{raw}	p_{Holm}	d_z
Biomechanics + PCA	Balanced accuracy	−0.0134	[−0.0279, −0.0004]	0.1282	0.2575	−0.565
Biomechanics + PCA	Macro-F1	−0.0161	[−0.0309, −0.0018]	0.0858	0.2575	−0.654
Biomechanics + PCA	ROC AUC	−0.0044	[−0.0099, 0.0013]	0.2754	0.2754	−0.458
Biomechanics + INR	Balanced accuracy	−0.0304	[−0.0426, −0.0178]	0.0077	0.0307	−1.446
Biomechanics + INR	Macro-F1	−0.0347	[−0.0485, −0.0205]	0.0020	0.0234	−1.469
Biomechanics + INR	ROC AUC	−0.0134	[−0.0175, −0.0091]	0.0039	0.0234	−1.887
INR only	Balanced accuracy	−0.1498	[−0.1737, −0.1260]	0.0020	0.0234	−3.666
Raw-waveform PCA only	Balanced accuracy	−0.2456	[−0.2744, −0.2147]	0.0020	0.0234	−4.826

Table C.22
Spearman correlations between participant-level INR reconstruction quality and absolute chronological age-prediction error.

Prediction model	Reconstruction metric	Spearman ρ	p
Biomechanics + INR	Mean reconstruction R^2	−0.050	0.676
Biomechanics + INR	Mean reconstruction RMSE	−0.043	0.717
Biomechanics + INR	Mean reconstruction MAE	−0.084	0.481
INR only	Mean reconstruction R^2	−0.142	0.230
INR only	Mean reconstruction RMSE	0.136	0.250
INR only	Mean reconstruction MAE	0.112	0.345
PCA + INR	Mean reconstruction R^2	−0.132	0.264
PCA + INR	Mean reconstruction RMSE	0.129	0.277
PCA + INR	Mean reconstruction MAE	0.110	0.356

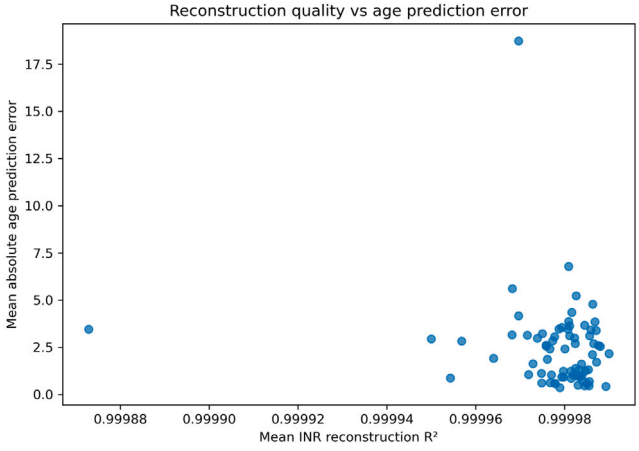

Fig. C.4. Relationship between participant-level reconstruction quality from the compact Fourier-MLP model and absolute chronological age-prediction error. Reconstruction R^2 values were strongly concentrated near one, resulting in substantial range restriction along the horizontal axis. The weak visual association is consistent with the non-significant Spearman correlations in Table C.22, but should be interpreted descriptively because of the restricted reconstruction range.

C.5. Morphological preservation and INR convergence diagnostics

To address whether INR fitting may smooth or attenuate individual-specific waveform characteristics, we compared original and reconstructed waveforms using amplitude-, timing-, and derivative-based descriptors. Global amplitude-related quantities were generally preserved with small absolute errors, particularly for ankle and hip moment trajectories. However, local dynamical quantities, including derivative energy and extrema count, were more sensitive to reconstruction. This indicates that INR reconstructions can preserve the global waveform envelope while still smoothing or altering local morphological details. This effect was most apparent for knee flexion–extension trajectories in terms of derivative-energy error, whereas the largest extrema-count discrepancies occurred for ankle moment trajectories.

The reconstruction values in this subsection were obtained with the compact Fourier-MLP model Appendix A.3, not the shallow-MLP protocol of Table C.19. The analysis covered the same five waveform classes and two sides for 73 participants, giving $73 \times 5 \times 2 = 730$ fitted curves.

Convergence diagnostics for the compact Fourier-MLP model are summarized in Table C.24. The model achieved a mean per-curve reconstruction R^2 of 0.999978 across 730 fitted curves, with a median value close to one. Final and minimum training losses were also consistently low. Most fits reached the maximum epoch limit rather than satisfying the early-stopping criterion.

These results show that the limited downstream utility of the Fourier-MLP-derived features cannot be attributed primarily to an inability to approximate the sampled waveforms.

Table C.23
Selected morphology-preservation errors between original and INR-reconstructed waveforms. Values are mean absolute errors across subjects.

Signal	Side	ROM error	Peak-timing error t_{max}	Derivative-energy error	Extrema-count error
Ankle dorsiflexion/plantarflexion moment	L	0.0015	0.0016	0.0195	9.12
Ankle dorsiflexion/plantarflexion moment	R	0.0017	0.0012	0.0214	10.36
Hip adduction/abduction moment	L	0.0017	0.0056	0.0171	1.78
Hip adduction/abduction moment	R	0.0015	0.0010	0.0197	1.77
Hamstring length	L	0.0002	0.0015	0.0003	1.05
Hamstring length	R	0.0002	0.0286	0.0003	1.23
Knee flexion–extension	L	0.0918	0.0018	48.1380	1.07
Knee flexion–extension	R	0.0738	0.0018	39.1226	1.14
Pelvic tilt	L	0.0066	0.0011	0.4633	0.25
Pelvic tilt	R	0.0091	0.0011	0.6583	0.78

Table C.24
In-sample convergence and reconstruction diagnostics for the compact Fourier-MLP model used in the matched PCA–INR benchmark. The model used 48–16 ReLU-activated hidden layers, per-curve target standardization, a maximum of 350 epochs, and an early-stopping patience of 30 epochs.

Architecture	Curves	Mean epochs	Final loss	Min. loss	Recon. R^2	Recon. RMSE
Compact Fourier MLP (48–16)	730	349.94	2.4×10^{-5}	2.2×10^{-5}	0.999978	0.014844

Data availability

The data that has been used is confidential.

References

[1] K.N. Dibbern, M.G. Krzak, A. Olivas, M.V. Albert, J.J. Krzak, K.M. Kruger, Scoping review of machine learning techniques in marker-based clinical gait analysis, Bioengineering 12 (2025) 591, http://dx.doi.org/10.3390/bioengineering12050591.

[2] E. Halilaj, A. Rajagopal, M. Fiterau, J.L. Hicks, T.J. Hastie, S.L. Delp, Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities, J. Biomech. 81 (2018) 1–11, http://dx.doi.org/10.1016/j.jbiomech.2018.09.009.

[3] J. Pauk, J. Szymul, Differences in pediatric vertical ground reaction force between planovalgus and neutrally aligned feet, Acta Bioeng. Biomech. 16 (2014) 95–101.

[4] L.M. Alderson, A. Wright, R. Wren, R. Kerr, M. Sangeux, et al., Age-related gait standards for healthy children and young people: The GOS-ICH paediatric gait centiles, Arch. Dis. Child. 104 (2019) 757–763, http://dx.doi.org/10.1136/archdischild-2018-315481.

[5] M.H. Schwartz, A. Rozumalski, The gait deviation index: A new comprehensive index of gait pathology, Gait Posture 28 (2008) 351–357, http://dx.doi.org/10.1016/j.gaitpost.2008.05.001.

[6] M.M. Bach, A. Daffertshofer, N. Dominici, The development of mature gait patterns in children during walking and running, Eur. J. Appl. Physiol. 121 (2021) 1073–1085, http://dx.doi.org/10.1007/s00421-020-04592-2.

[7] J. Burdack, F. Horst, S. Giesselbach, I. Hassan, S. Daffner, W.I. Schöllhorn, Systematic comparison of the influence of different data preprocessing methods on the performance of gait classifications using machine learning, Front. Bioeng. Biotechnol. 8 (2020) 260, http://dx.doi.org/10.3389/fbioe.2020.00260.

[8] W. Liu, Q. Mei, P. Yu, Z. Gao, Q. Hu, G. Fekete, I. Bíró, Y. Gu, Biomechanical characteristics of the typically developing toddler gait: A narrative review, Children 9 (2022) 406, http://dx.doi.org/10.3390/children9030406.

[9] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Springer, New York, 2005.

[10] P.A. Federolf, K.A. Boyer, T.P. Andriacchi, Application of principal component analysis in clinical gait research: Identification of systematic differences between healthy and medial knee-osteoarthritic gait, J. Biomech. 46 (2013) 2173–2178, http://dx.doi.org/10.1016/j.jbiomech.2013.06.032.

[11] V. Ronca, R. Capotorto, S. Buonocore, F. Massa, G. Di Gironimo, R. De Amicis, S. Papa, L. Mattioli, E. Palermo, L. Di Donato, D. Freda, M. Pirozzi, A. Ferraro, G. Di Flumeri, A. Giorgi, G. Borghini, P. Aricò, A novel neurophysiological approach to evaluate the impact of virtual training on skills acquisition, J. Neural Eng. 23 (2026) 026032, http://dx.doi.org/10.1088/1741-2552/ae5a06.

[12] S. Scataglini, E. Abts, C. Van Bocxlaer, M. Van den Bussche, S. Meletani, S. Truijen, Accuracy, validity, and reliability of markerless camera-based 3D motion capture systems versus marker-based systems in gait analysis: A systematic review and meta-analysis, Sensors 24 (2024) 3686, http://dx.doi.org/10.3390/s24113686.

[13] V. Sitzmann, J.N.P. Martel, A.W. Bergman, D.B. Lindell, G. Wetzstein, Implicit neural representations with periodic activation functions, Adv. Neural Inf. Process. Syst. 33 (2020) 7462–7473.

[14] M. Tancik, P.P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J.T. Barron, R. Ng, Fourier features let networks learn high frequency functions in low dimensional domains, Adv. Neural Inf. Process. Syst. 33 (2020) 7537–7547.

[15] J. Lee, J. Tack, N. Lee, J. Shin, Meta-learning sparse implicit neural representations, Adv. Neural Inf. Process. Syst. 34 (2021) 11091–11103.

[16] B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng, NeRF: Representing scenes as neural radiance fields for view synthesis, in: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., CVPR, 2020, pp. 4055–4064, http://dx.doi.org/10.1109/CVPR42600.2020.00412.

[17] J.J. Park, P. Florence, J. Straub, R. Newcombe, S. Lovegrove, DeepSDF: Learning continuous signed distance functions for shape representation, in: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., CVPR, 2019, pp. 165–174, http://dx.doi.org/10.1109/CVPR.2019.00024.

[18] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, A. Geiger, Occupancy networks: Learning 3D reconstruction in function space, in: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., CVPR, 2019, pp. 4460–4470, http://dx.doi.org/10.1109/CVPR.2019.00460.

[19] L. Yariv, J. Gu, Y. Kasten, Y. Lipman, V. Koltun, Volume rendering of neural implicit surfaces, Adv. Neural Inf. Process. Syst. 34 (2021) 4805–4815.

[20] S. Peng, C. Geng, Y. Zhang, Y. Xu, Q. Wang, Q. Shuai, X. Zhou, H. Bao, Implicit neural representations with structured latent codes for human body modeling, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2023) 9895–9907, http://dx.doi.org/10.1109/TPAMI.2023.3246809.

[21] Y. Zhu, Y. Liu, Y. Zhang, D. Liang, Implicit neural representation for medical image reconstruction, Phys. Med. Biol. 70 (2025) 125001, http://dx.doi.org/10.1088/1361-6560/addfa5.

[22] R.T.Q. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, Neural ordinary differential equations, Adv. Neural Inf. Process. Syst. 31 (2018) 6571–6583.

[23] P. Kidger, J. Morrill, J. Foster, T. Lyons, Neural controlled differential equations for irregular time series, Adv. Neural Inf. Process. Syst. 33 (2020) 6696–6707.

[24] L. Xiang, Z. Gao, P. Yu, J. Fernandez, Y. Gu, R. Wang, E.M. Gutierrez-Farewik, Explainable artificial intelligence for gait analysis: Advances, pitfalls, and challenges, Front. Bioeng. Biotechnol. 13 (2025) 1671344, http://dx.doi.org/10.3389/fbioe.2025.1671344.

[25] J. Figueiredo, C.P. Carvalho, D. Gonçalves, A. Vilas-Boas, J. Vilas-Boas, Machine learning strategies for adult and pediatric gait classification: A systematic review, IEEE Rev. Biomed. Eng. 11 (2018) 233–251, http://dx.doi.org/10.1109/RBME.2018.2836314.

[26] J. Warmenhoven, N. Bargary, D. Liebl, G. Hooker, PCA of waveforms and functional PCA: A primer for biomechanics, J. Biomech. 116 (2021) 110106, http://dx.doi.org/10.1016/j.jbiomech.2020.110106.

[27] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1798–1828, http://dx.doi.org/10.1109/TPAMI.2013.50.

[28] T.S. Winner, M.C. Rosenberg, G.J. Berman, T.M. Kesar, L.H. Ting, Gait signature changes with walking speed are similar among able-bodied young adults despite persistent individual-specific differences, Sci. Rep. 14 (2024) 19730, http://dx.doi.org/10.1038/s41598-024-70787-8.

[29] G. Bovi, M. Rabuffetti, P. Mazzoleni, M. Ferrarin, A multiple-task gait analysis approach: Kinematic, kinetic and EMG reference data for healthy young and adult subjects, Gait Posture 33 (2011) 6–13, http://dx.doi.org/10.1016/j.gaitpost.2010.08.009.

[30] E. Dupont, H. Kim, S.M.A. Eslami, D.J. Rezende, D. Rosenbaum, From data to functa: Your data point is a function and you can treat it like one, in: Proc. Int. Conf. Mach. Learn., ICML, 2022, pp. 5694–5713.

[31] K. Liu, F. Liu, H. Wang, N. Ma, J. Bu, B. Han, Partition speeds up learning implicit neural representations based on exponential-increase hypothesis, in: 2023 IEEE/CVF International Conference on Computer Vision, ICCV, Paris, France, 2023, pp. 5451–5460, http://dx.doi.org/10.1109/ICCV51070.2023.00504.