

## The 1st AI Children Challenge

Boyi Li<sup>3</sup>, Yifan Shen<sup>3</sup>, Houze Yang<sup>3</sup>, Xu Cao<sup>2,3</sup>, Guojun Yun<sup>1</sup>, Li Gao<sup>1</sup>,  
Turong Chen<sup>1</sup>, Long Xu<sup>4</sup>, Jianguo Cao<sup>1,2</sup>, Meihuan Huang<sup>1,5,\*</sup>

<sup>1</sup>Shenzhen Children's Hospital   <sup>2</sup>PediaMed AI   <sup>3</sup>University of Illinois Urbana-Champaign  
<sup>4</sup>Xiangya Bo'ai Rehabilitation Hospital   <sup>5</sup>The Hong Kong Polytechnic University

boyil2@illinois.edu; meihuan.huang@connect.polyu.hk

### Abstract

*The First AI Children Challenge aims to advance real-world applications of computer vision and AI in child healthcare, child education, and pediatrics. The 2026 CV4CHL edition featured the first track in this domain: Children Gait Visual Analysis. The main goal of Children Gait Visual Analysis is the fine-grained analysis of children's gait behaviors from keypoint sequences. This is still a big challenge for human action recognition. Experienced medical doctors can distinguish these subtle nuances, but none of the people test AI models in this domain. To bridge this gap, we introduce thousands of 2D children keypoint sequences walking around videos across various age groups of children (3-16 years old). There is a significant opportunity for batch analysis of these videos to provide clinically relevant insights into medical diagnosis. The Challenge will be launched with two problem tracks: **Edinburgh Visual Gait Score (EVGS) Scoring and Classification of Gait Patterns in Bilateral Spastic Cerebral Palsy**. Each track is chosen in consultation with board-certified pediatricians based on the value of potential solutions. With the first available dataset for such tasks and ground truth for each track, the challenge enabled participants to evaluate their solutions. Final rankings will be revealed after the competition concludes, fostering reproducibility and mitigating overfitting.*

### 1. Introduction

Quantitative gait analysis is a crucial component of clinical diagnostics and physical rehabilitation for movement disorders [3–5, 19, 21]. Understanding pediatric gait patterns through various developmental milestones is crucial for diagnosing and managing several developmental disorders [1,

\*Corresponding author

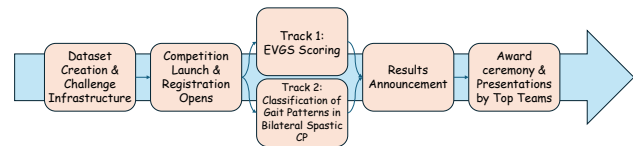


Figure 1. **The Life Cycle of the First AI Children Challenge.** The challenge spanned the entire process, from preparation and launch to the award ceremony. Participants may enter both Track 1 and Track 2 simultaneously.

[2, 17, 23]. The early identification of gait abnormalities in conditions such as cerebral palsy primarily depends on subjective visual assessments conducted by a pediatrician during standard office visits [10, 15]. Although research using 3D gait analysis and Inertial Measurement Units (IMUs) has successfully identified quantitative warning signs for gait deviations in cerebral palsy [8, 9], implementing these technologies in clinical practice encounters significant challenges. For example, the cost of 3D gait analysis device is high and children may have uncooperative engagement. Besides, existing computer vision methods for gait analysis have largely focused on healthy adult subjects [7, 16], relying on the assumption that walking is a mature, stable, and highly rhythmic process [12]. Such adult-centric foundation models fail to account for the high entropy, significant intra-class variance, and inconsistent motion patterns inherent in the developing motor system [20, 22].

To explore potential solutions to these issues in greater depth, we define a new challenge for the computer vision community: *the fine-grained analysis of children's gait from video, targeting clinically-relevant characterizations of children's gait quality.*, and will host the first edition of the AI Children Challenge in which we introduce the Children Gait Visual Analysis track of this new domain, aiming to drive advancements in computer vision and AI for real-world, high-impact domains. Figure 1 illustrates the life cycle of the challenge. Preparation for the Challenge began in

Table 1. **17 Scoring Items of EVGS.** The Edinburgh Visual Gait Score (EVGS) is a clinical tool developed to visually assess gait deviations in ambulatory children using coronal and sagittal video recordings. It evaluates 17 observational scoring items for each limb that are graded on a three-point ordinal scale.

| No. | Gait Pattern                           | Abbr. | Score Scale | View     |
|-----|--|-------|-------------|----------|
| 1   | Initial Contact in Stance              | IC    | 0, 1        | Sagittal |
| 2   | Heel Lift in Stance                    | HL    | 0, 1        | Sagittal |
| 3   | Max Ankle Dorsiflexion in Stance       | SAD   | 0, 1        | Sagittal |
| 4   | Hind-foot Varus/Valgus in Stance       | HVV   | 0, 1        | Coronal  |
| 5   | Foot Rotation in Stance                | FRT   | 0, 1        | Coronal  |
| 6   | Foot Clearance in Swing                | FCL   | 0, 1        | Sagittal |
| 7   | Max Ankle Dorsiflexion in Swing        | WAD   | 0, 1        | Sagittal |
| 8   | Knee Progression Angle in Mid-Stance   | KPA   | 0, 1        | Coronal  |
| 9   | Peak Knee Extension in Stance          | KEX   | 0, 1        | Sagittal |
| 10  | Knee Position in Terminal Swing        | KPS   | 0, 1        | Sagittal |
| 11  | Peak Knee Flexion in Swing             | KFX   | 0, 1        | Sagittal |
| 12  | Peak Hip Extension in Stance           | HEX   | 0, 1        | Sagittal |
| 13  | Peak Hip Flexion during Swing          | HFX   | 0, 1        | Sagittal |
| 14  | Pelvic Obliquity at Mid-Stance         | POB   | 0, 1        | Coronal  |
| 15  | Pelvic Rotation at Mid-Stance          | PRT   | 0, 1        | Coronal  |
| 16  | Peak Sagittal Trunk Position in Stance | TSG   | 0, 1        | Sagittal |
| 17  | Maximum Trunk Lateral Shift            | TLT   | 0, 1        | Coronal  |

January 2026, and the Challenge will end in the CVPR 2026 Workshop on Computer Vision for Children (CV4CHL).

To support this Challenge, we introduce the Children Gait Pose Sequence (CGPS) dataset, which features thousands of 60 FPS 2D human keypoint sequences capturing the natural walking behaviors of children across various age groups (2 to 16 years old), as detailed in Section 3. With the first available dataset for such tasks and a comprehensive ground truth for each track, the Challenge enables participating teams to effectively evaluate their solutions.

The Challenge will be launched with two specific problem tracks: **Edinburgh Visual Gait Score (EVGS) Scoring** and **classification of Gait Patterns in Bilateral Spastic Cerebral Palsy**. Each track is chosen in consultation with board-certified pediatricians based on the clinical value of potential solutions, and they are all interconnected in some way [23].

This paper presents a comprehensive summary of the preparation and results of the First AI Children Challenge. The following sections describe the challenge setup Section 2, dataset preparation Section 3, evaluation methodology Section 4, results of participant submissions Section 5, and conclude with a discussion of the findings and future research directions Section 6.

## 2. Challenge Setup

The First AI Children Challenge will be held on Kaggle [14]. It allows participants to compete in one or all of the following two tracks. The Challenge follows a structured timeline, with the Challenge launching and registration opening on March 23, 2026. Participants are required to submit their challenge results by April 23, 2026. For

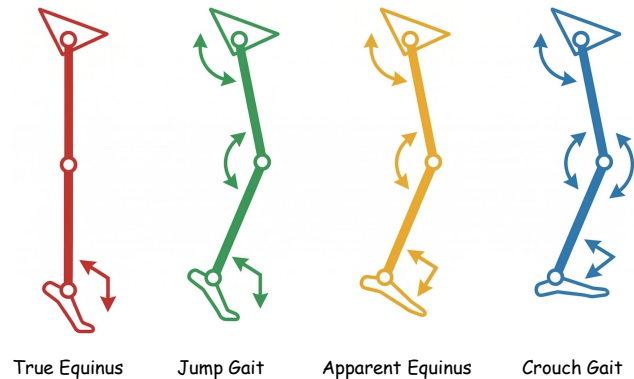


Figure 2. **The Summary of the Classification of Gait Patterns in Bilateral Spastic Cerebral Palsy.** From left to right are four different gait patterns in bilateral spastic CP: **True Equinus (red):** Defined by hip and knee extension (sometimes knee recurvatum) and an equinus foot position. **Jump Gait (green):** Characterized by hip and knee flexion with an equinus foot position, typically associated with anterior pelvic tilt and lumbar lordosis. **Apparent Equinus (yellow):** Features increased hip and knee flexion and a decreased equinus ankle angle, leading to increased dorsiflexion. **Crouch Gait (blue):** Marked by excessive hip and knee flexion, scissoring, and excessive dorsiflexion.

teams intending to publish accompanying papers, the workshop paper (non-proceeding track) submission deadline was set for May 15, 2025. All deadlines are at 11:59 PM (Anywhere on Earth Time) of the corresponding day unless otherwise stated.

To promote transparency and reproducibility, teams competing for top rankings were required to publicly release their code. This policy ensured that leaderboard results could be independently verified and contributed to the broader research community

**Track 1 - EVGS Scoring.** As shown in Table 1, Edinburgh Visual Gait Score (EVGS) [18] is a kind of observational scale which is critical for diagnosing neurodevelopmental disorders such as Cerebral Palsy (CP). Therefore, designing dedicated pediatric modeling techniques and developing a model capable of automatic scoring to provide pediatricians with a reference is essential for clinical diagnosis. In this track, participants are tasked with developing algorithms capable of conducting fine-grained kinematic analysis to accurately predict EVGS metrics directly from provided children’s 2D keypoint sequences. Evaluation of Track 1 is based on distinguishing subtle kinematic differences, measured by Accuracy and RMSE, which tells the accuracy of the predictions and the difference between the predicted score and the ground truth for each item.

**Track 2 - Classification of Gait Patterns in Bilateral Spastic Cerebral Palsy.** Although Gait classification systems for children with bilateral lower limb spasticity have

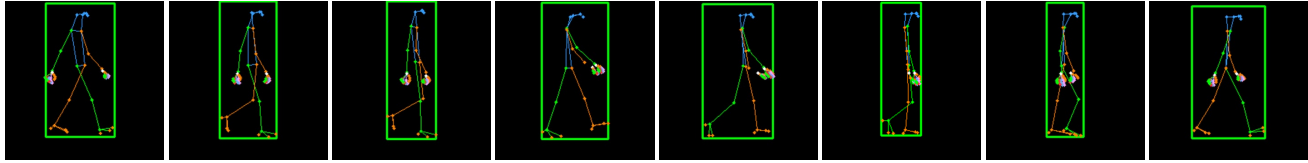


Figure 3. **Dataset Examples Visualization.** The CGPS dataset provides keypoint sequences and bounding boxes per frame. We use SAM3 [6] to perform instance segmentation and Sapiens-2B [13] to perform pose estimation to obtain bounding boxes and keypoints. We will release the keypoint sequences for the participants.

been developed, there are limitations to their clinical relevance and applicability in orthotic prescription. Studies have shown low levels of validity and reliability for these classification systems [11], emphasizing the need for comprehensive approaches that address the range and magnitude of gait deviations in children with bilateral spastic CP. A summary of the classification of gait patterns in bilateral spastic CP is shown in Figure 2. As a track running parallel to Track 1, this track focuses on dealing with this issue, which challenges teams to classify specific gait patterns of bilateral spastic cerebral palsy by recognizing subtle, pathological variations in children’s gait behaviors. Evaluation of Track 2 is based on the clinical diagnostic performance, measured by the accuracy rate and the F1 score.

### 3. Children Gait Pose Sequence Dataset

We introduce the Children Gait Pose Sequence (CGPS) dataset to support this Challenge. Before the data annotation and model design, the IRB approval is obtained from affiliated hospitals. The dataset contains 2D keypoints annotation per frame across a total number of 339,236 frames (1,185 videos) of size  $1920 \times 1080$  (2K) with 17 EVGS [18] sub-item annotations per limb and the patient-level diagnosis results of Cerebral Palsy (CP) motor types (Unilateral or Bilateral) and corresponding gait patterns. The videos were recorded at Shenzhen Children’s Hospital using smartphone cameras and action cameras, positioned simultaneously to capture sagittal and coronal views. We then use SAM3 [6] to detect bounding boxes and Sapiens-2B [13] to detect 2D keypoint sequences, where all annotations are manually adjusted by human annotators. Figure 3 shows examples of videos and annotations. In total, 110 patients participated in the data collection. In some cases, multiple recordings per view were available. The dataset provides detailed annotations per frame and per video, including the subject’s body bounding box (detected with SAM 3 [6] and manually selected by human annotators), the 2D human keypoints (detected with Sapiens-2B [13] and manually adjusted by human annotators), the EVGS sub-items, CP motor types and corresponding gait patterns (annotated by an experienced pediatricians in the author team and reviewed by a senior pediatrician author with 40 years of clinical experi-

ence). Figure 4 shows the EVGS scores distribution, CP motor types distribution, and bilateral spastic CP gait patterns distribution among all patients.

To capture the dataset, a primary camera is positioned at the terminus of an 8-meter walkway to record the coronal (frontal and posterior) view. A secondary camera is oriented orthogonally, facing the center of the walkway, to capture the sagittal (lateral) view. This lateral camera is positioned at a sufficient distance to ensure its field of view encompasses the middle four meters of the trial space. This specific distance is calibrated to guarantee the capture of 2-3 complete gait cycles (strides) per subject. Then the videos are further processed to obtain bounding boxes and key point sequences per frame. Table 1 details the 17 fine-grained gait parameters annotated in the CGPS dataset, which strictly adhere to the EVGS reference guide to ensure high clinical validity. Figure 2 details the four gait patterns in bilateral spastic CP. While the standard EVGS protocol employs a three-point ordinal scale (0: normal, 1: moderate deviation, 2: severe deviation) for each parameter, the natural distribution of pediatric gait pathologies inherently results in a severe class imbalance, particularly for the most extreme deviations (score 2). To mitigate this imbalance and establish a robust computational benchmark, we binarize the assessment by merging scores 1 and 2. Consequently, the prediction for each fine-grained gait parameter is formulated as a binary classification task (see Figure 4) discriminating between “typical” and “atypical” gait patterns.

The age and gender distribution of the patients are shown in Figure 4. The age histogram is overlaid with a fitted normal distribution curve to illustrate the central tendency of the patient demographics. The participants span a critical developmental window from 2.6 to 16.6 years old, with a mean age of  $\mu = 8.40$ ,  $\sigma = 3.14$  years, which is highly clinically relevant and highlights a primary focus on the pediatric population. It also details the gender composition of the dataset. The patients consist of 58.2% male and 41.8% female subjects. This nearly balanced gender distribution helps to prevent the model from learning shortcuts for specific genders, ensuring sufficient diversity and fairness for training and evaluating models for visual gait analysis.

To prevent identity leakage and shortcut bias, we imple-

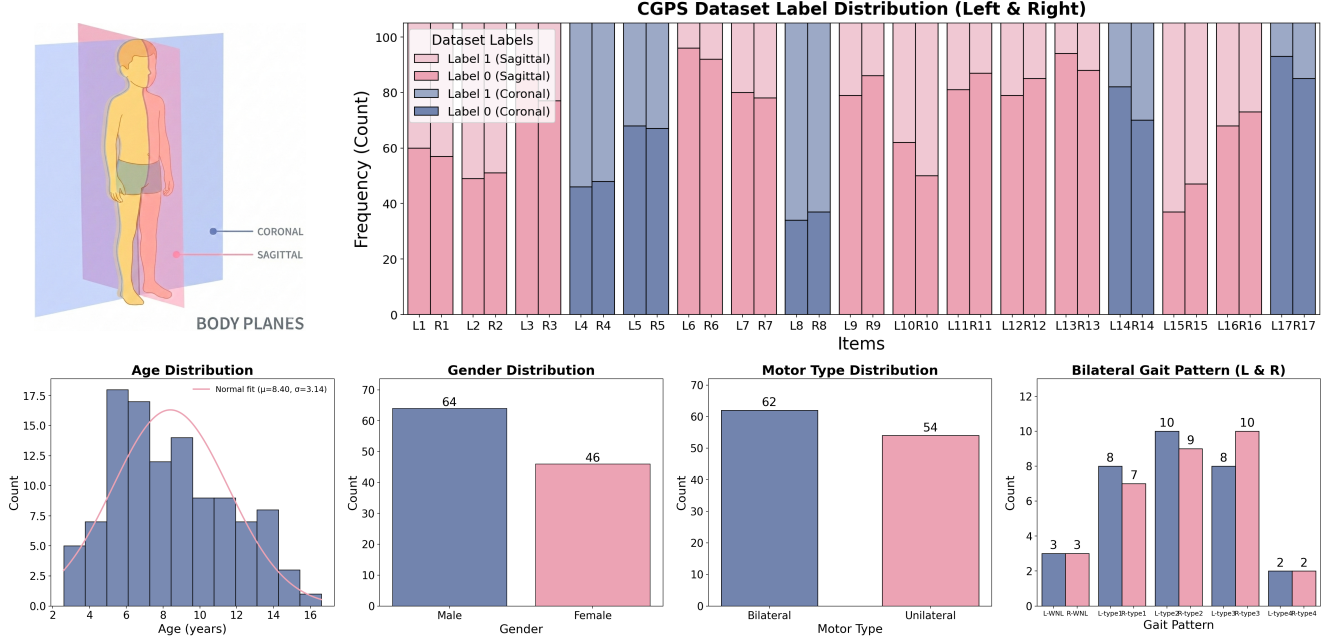


Figure 4. **The Overview of our CGPS Dataset.** The CGPS dataset comprises 110 pediatric patients, with each evaluation covering a total of  $2 \times 17$  EVGS scoring items (17 items per limb) across two body planes, a CP motor type, and a corresponding gait pattern. *Top Left:* The child’s body planes. *Top Right:* The label distribution of the CGPS dataset. *Bottom:* From left to right: age distribution, gender distribution, CP motor type distribution, and bilateral spastic CP gait patterns distribution. Overall, the dataset maintains a nearly balanced profile, providing the necessary diversity for training or evaluation in visual gait analysis.

ment a strict object-level data split based on the unique **Patient ID** in the experiment. For different tracks, the test set is randomly selected, and we ensure all annotations (scores distribution in Track 1 and gait patterns distribution in Track 2) are sampled nearly balanced. All sequences, corresponding annotations, and derived gait cycles for any given patient belong exclusively to either the training or the test set. For different tracks, we select different sets of patients as test sets and will only release sequences corresponding to the track and specific annotations.

## 4. Evaluation Protocols

The First AI Children Challenge featured two tracks spanning EVGS scoring and gait patterns classification. The datasets, task objectives, submission formats, and evaluation metrics are summarized below.

### 4.1. Track 1 Evaluation

Track 1 challenges teams to conduct fine-grained kinematic analysis and accurately predict Edinburgh Visual Gait Score (EVGS) metrics from the provided multi-view annotations. The core of this evaluation focuses on distinguishing subtle kinematic differences, which are measured by the discrepancy between the predicted score and the clinical ground truth for each EVGS item. The dataset covers the entire CGPS dataset, carefully selecting the evaluation set to ensure that the number of positive and negative samples is

nearly equal, ensuring the accuracy of the test. The ratio of the training set to the evaluation set is 94:16.

Each patient provides multi-view JSON files, including 2D keypoint sequences and bounding boxes per frame, and the EVGS scores per patient. Participants are required to design algorithms and models to track the correct patients, analyze the gait features, and score the EVGS items based on the criteria automatically.

In this track, we use Accuracy and Root Mean Square Error (RMSE) to comprehensively evaluate the predictive ability of the fine-grained kinematic analysis and EVGS scoring of the model. Specifically, for each patient in the evaluation set, we will calculate the overall accuracy:

$$\text{Accuracy}_{\text{track1}} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}[y_i = \hat{y}_i] \quad (1)$$

where  $N$  is the total number of scoring items across all patients in the evaluation set of Track 1,  $\hat{y}_i$  is the predicted EVGS score, and  $y_i$  is the clinical ground truth,  $\hat{y}_i, y_i \in \{0, 1\}$ .

Besides, we calculate the patient-level error. The error between the ground truth and the predicted values is calculated using the Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (p_i - g_i)^2} \quad (2)$$

Listing 1. **Required Submission Format.** Participants should fill in the cells under the requirements, and leave the blank cells with -1.

```

ID, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11, L12, L13, L14, L15, L16, L17, R1, R2, R3, R4, R5, R6, R7, R8, R9
, R10, R11, R12, R13, R14, R15, R16, R17, Total, Left_gait_subtype, Right_gait_subtype
track1-4, ...
...
track2-4, ...
...

```

where  $M$  is the number of sampled patients in the evaluation set of Track 1,  $p_i = \sum_j \hat{y}_j$  is the predicted total EVGS score for a particular patient  $i$ , and  $g_i = \sum_j y_j$  is the clinical ground truth for the particular patient  $i$ .

To ensure that a smaller gap translates to a higher final score, the evaluation score for Track 1, denoted as  $S_1$ , uses the Accuracy and normalized RMSE (NRMSE) with equal weights to scale the results relative to all participating teams:

$$\text{NRMSE} = \frac{\text{RMSE}}{34} \quad (3)$$

$$S_1 = \frac{\text{Accuracy}_{\text{Track1}} + 1 - \text{NRMSE}}{2} \quad (4)$$

This scoring mechanism effectively rewards models that most closely align with the precise ground-truth measurements.

## 4.2. Track 2 Evaluation

Track 2 evaluates the model’s clinical diagnostic performance in accurately classifying specific gait patterns of bilateral spastic CP. This task challenges algorithms to recognize subtle, pathological variations in children’s gait behaviors that dictate different bilateral lower limb spasticity classifications. We select a subset from the CGPS dataset consisting exclusively of patients with bilateral spastic CP. The evaluation set is also carefully selected to ensure that the number of five gait pattern samples is nearly equal. The ratio of the training set to the evaluation set is 22:9.

Similar to Track 1, each patient provides multi-view JSON files, including 2D keypoint sequences and bounding boxes per frame, and detailed gait patterns per patient. Participants are required to design algorithms and models to track the correct patients, analyze the gait features, and score the EVGS items based on the criteria automatically.

To rank participant submissions, we use Accuracy and F1 Score, designed to balance recall and precision. This ensures that models do not over-predict specific types (leading to low precision) nor under-predict specific types (resulting in low recall). There are several variations of the F1 score (macro-averaged per sample, macro-averaged per class, and micro-averaged). For this challenge, we use the macro-averaged F1 score per sample, which is particularly suited

for multi-label classification problems. This approach ensures a fair evaluation of each image independently, rather than being influenced by class distribution.

$$\text{F1}_k = \frac{2P_k R_k}{P_k + R_k} \quad (5)$$

where for each gait pattern in bilateral spastic CP  $k \in \{\text{TE, JG, AE, CG, WNL}\}$ , we consider it as the positive class, and the other four categories as the negative classes. Therefore, the macro-averaged F1 score is:

$$\text{F1}_{\text{macro}} = \frac{1}{5} \sum_{k=1}^5 \text{F1}_k \quad (6)$$

We also calculate the overall accuracy for each patient in the evaluation set:

$$\text{Accuracy}_{\text{track2}} = \frac{1}{M} \sum_{k=1}^5 \text{TP}_k \quad (7)$$

where  $M$  is the number of sampled patients in the evaluation set of Track 2. Notice that the evaluation sets are different in Track 1 and Track 2.

Denote the evaluation score of Track 2 as  $S_2$ , which is calculated by aggregating the macro-averaged F1 score and Accuracy with equal weights, reflecting holistic clinical utility and robustness:

$$S_2 = \frac{\text{Accuracy}_{\text{Track2}} + \text{F1}_{\text{macro}}}{2} \quad (8)$$

## 4.3. Evaluation System

The First AI Children Challenge used a centralized evaluation portal via Kaggle [14]. Participants created an evaluation account with an email and password and verified their email to activate the account. Once registered, teams could submit their predictions.

We just set up one overall leaderboard that combines the scores of two tracks. During the Challenge, teams could see a leaderboard with the best results from each team. The rankings are based on the total scores from both tracks, and the scores of the two tracks each account for 50% to determine the Final Score  $S$ :

$$S = \frac{S_1 + S_2}{2} \quad (9)$$

Table 2. **Summary of the Final Score  $S$ .** We compile the participation statistics as of April 8 at 11.59 PM (Anywhere on Earth Time).

| Rank | Team           | Score   | Entries |
|------|----------------|---------|---------|
| 1    | seantangth     | 0.88294 | 77      |
| 2    | ylmoonie       | 0.70504 | 30      |
| 3    | Spiritmilk     | 0.60214 | 30      |
| 4    | M45t3rL4       | 0.59322 | 34      |
| 5    | CCE            | 0.52321 | 12      |
| 6    | D44111090      | 0.51789 | 22      |
| 7    | kunli-cs       | 0.51278 | 6       |
| 8    | zuw96d3f51ts   | 0.49795 | 3       |
| 9    | POU            | 0.49689 | 3       |
| 10   | Yutong He 2025 | 0.49507 | 17      |
| 11   | young          | 0.48576 | 4       |
| 12   | Zihan Ji       | 0.46069 | 8       |
| 13   | Xiaochuan Guo  | 0.40250 | 11      |
| 14   | MR-CAS         | 0.39423 | 5       |
| 15   | AIThings       | 0.39375 | 17      |

The submission file should be in CSV format, which contains both track results. The file should contain a header. Listing 1 shows the required format for submission. Participants should use `track1-` and `track2-` as prefixes, and then add the patient ID to distinguish between Track 1 and 2. For cells that do not belong to the current track or the track that they do not participate in, participants are required to enter `-1`. The results of Track 1 should be listed first, followed by Track 2, and the IDs should be sorted in ascending order in the first column. For Track 1, the cell should be filled with `0` or `1`, indicating negative or positive. For Track 2, the cell should be filled with one in **type1**, **type2**, **type3**, **type4**, **WNL**.

To qualify for awards via the leaderboard, teams are required to:

- Cannot enter or submit from multiple accounts.
- Publicly release reproducible code, train models before the deadline to verify the submission, promote transparency and reproducibility.
- Write and submit a technical report to the CV4CHL Workshop non-proceeding track.

Submission limits were set as follows:

- Up to 10 submissions per day (Anywhere on Earth Time).
- Only select 1 final submission for judging.

After the Challenge ends, we will re-evaluate the submissions locally using the released code and model from teams. Any team that has not made its code public, or any team whose results differ significantly from the local evaluation, will be deemed non-compliant and will forfeit its eligibility for awards.

## 5. Challenge Results

Table 2 provides a summary of the leaderboard results for the Final Score  $S$ . As of April 8, 11:59 PM (Anywhere on Earth), a total of 15 teams participated in our Challenge, with the Final Score ranging from 39% to 88%.

## 6. Discussion and Conclusion

The First AI Children Challenge marks a significant milestone in advancing applied computer vision across child healthcare, child education, and pediatrics. With newly proposed datasets and novel benchmarks, the Challenge pushed the boundaries of fine-grained children’s gait analysis.

Track 1 introduces clinical annotations of professional pediatric doctors’ EVGS scores for the CGPS dataset. This is an asset of critical importance for pediatric diagnostics. A key research opportunity remains in automated severity assessment, especially in handling subtle pathological cues and high inter-patient variability in pediatric cases. Future iterations may further challenge generalization with increased clinical complexity, such as varied imaging conditions or diverse anatomical appearances.

Track 2 shifts the analytical focus toward bilateral spastic, a specific motor type of Cerebral Palsy, by systematically investigating its four distinct sub-gait patterns. A key research opportunity remains in the robust differentiation of transitional or borderline gait patterns, especially when handling noisy data and atypical clinical presentations. Future iterations may further challenge generalization by expanding to other CP motor types or incorporating longitudinal data to evaluate gait evolution over time.

The First AI Children Challenge highlights a strong trend toward clinically-guided feature extraction, fine-grained kinematic modeling, and diagnostic interpretability. As AI systems move toward deployment in real-world clinical workflows, the continued convergence of computer vision, biomechanics, and medical expertise will drive innovation in precision pediatric healthcare.

## 7. Acknowledgement

The CGPS dataset is developed through extensive data curation efforts enabled by close collaboration. Key contributors include Shenzhen Children’s Hospital and PediaMed AI, alongside academic partners such as the University of Illinois Urbana-Champaign and The Hong Kong Polytechnic University. The Challenge also benefited from thorough review efforts by researchers. We would like to thank the gait research community for their previous work. We thank the broader research community for their participation, which helps establish the AI Children Challenge as a leading benchmark in pediatric healthcare. The contributions of participants, reviewers, and collaborators are essential to the success of the Challenge.

## References

- [1] Stéphane Armand, Geraldo Decoulon, and Alice Bonnefoy-Mazure. Gait analysis in children with cerebral palsy. *EFORT open reviews*, 1(12):448–460, 2016. 1
- [2] Richard Baker. Gait analysis methods in rehabilitation. *Journal of neuroengineering and rehabilitation*, 3(1):4, 2006. 1
- [3] Richard Baker, Alberto Esquenazi, Maria Grazia Benedetti, Kaat Desloovere, et al. Gait analysis: clinical facts. *Eur. J. Phys. Rehabil. Med*, 52(4):560–574, 2016. 1
- [4] Nawel Ben Chaabane, Pierre-Henri Conze, Mathieu Lempereur, Gwenolé Quellec, Olivier Rémy-Néris, Sylvain Brochard, Béatrice Cochener, and Mathieu Lamard. Quantitative gait analysis and prediction using artificial intelligence for patients with gait disorders. *Scientific Reports*, 13(1):23099, 2023.
- [5] Mirjam Bonanno, Alessandro Marco De Nunzio, Angelo Quartarone, Annalisa Miliati, Francesco Petralito, and Rocco Salvatore Calabrò. Gait analysis in neurorehabilitation: from research to clinical practice. *Bioengineering*, 10(7):785, 2023. 1
- [6] Nicolas Carion, Laura Gustafson, Yuan-Ting Hu, Shoubhik Debnath, Ronghang Hu, Didac Suris, Chaitanya Ryal, Kalyan Vasudev Alwala, Haitham Khedr, Andrew Huang, et al. Sam 3: Segment anything with concepts. *arXiv preprint arXiv:2511.16719*, 2025. 3
- [7] Biao Chen, Chaoyang Chen, Jie Hu, Zain Sayeed, Jin Qi, Hussein F Darwiche, Bryan E Little, Shenna Lou, Muhammad Darwish, Christopher Foote, et al. Computer vision and machine learning-based gait pattern recognition for flat fall prediction. *Sensors*, 22(20):7960, 2022. 1
- [8] Robert E Cook, Ingo Schneider, M Elizabeth Hazlewood, Susan J Hillman, and James E Robb. Gait analysis alters decision-making in cerebral palsy. *Journal of pediatric orthopaedics*, 23(3):292–295, 2003. 1
- [9] Peter A DeLuca, Roy B Davis, Sylvia Öunpuu, Sally Rose, and Robert Sirkin. Alterations in surgical decision making in patients with cerebral palsy based on three-dimensional gait analysis. *Journal of Pediatric Orthopaedics*, 17(5):608–614, 1997. 1
- [10] Wendy E Dickens and Michael F Smith. Validation of a visual gait assessment scale for children with hemiplegic cerebral palsy. *Gait & posture*, 23(1):78–82, 2006. 1
- [11] Fiona Dobson, Meg E Morris, Richard Baker, and H Kerr Graham. Gait classification in children with cerebral palsy: a systematic review. *Gait & posture*, 25(1):140–152, 2007. 3
- [12] Chao Fan, Junhao Liang, Chuanfu Shen, Saihui Hou, Yongzhen Huang, and Shiqi Yu. Opengait: Revisiting gait recognition towards better practicality. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9707–9716, 2023. 1
- [13] Rawal Khirodkar, Timur Bagautdinov, Julieta Martinez, Su Zhaoen, Austin James, Peter Selednik, Stuart Anderson, and Shunsuke Saito. Sapiens: Foundation for human vision models. In *European Conference on Computer Vision*, pages 206–228. Springer, 2024. 3
- [14] Boyi Li, Yifan Shen, Houze Yang, Xu Cao, Guojun Yun, Li Gao, Turong Chen, Long Xu, Jianguo Cao, and Meihuan Huang. [cvpr 2026] the first ai for children challenge. <https://kaggle.com/competitions/cvpr-2026-the-first-ai-children-challenge>, 2026. Kaggle. 2, 5
- [15] Karel GB Maathuis, Cees P Van Der Schans, Andries Van Iperen, Hans S Rietman, and Jan HB Geertzen. Gait in children with cerebral palsy: observer reliability of physician rating scale and edinburgh visual gait analysis interval testing scale. *Journal of Pediatric Orthopaedics*, 25(3):268–272, 2005. 1
- [16] Rahm Ranjan, David Ahmedt-Aristizabal, Mohammad Ali Armin, and Juno Kim. Computer vision for clinical gait analysis: A gait abnormality video dataset. *IEEE Access*, 13:45321–45339, 2025. 1
- [17] Chandrasekar Rathinam, Andrew Bateman, Janet Peirson, and Jane Skinner. Observational gait assessment tools in paediatrics—a systematic review. *Gait & posture*, 40(2):279–285, 2014. 1
- [18] Heather S Read, M Elizabeth Hazlewood, Susan J Hillman, Robin J Prescott, and James E Robb. Edinburgh visual gait score for use in cerebral palsy. *Journal of pediatric orthopaedics*, 23(3):296–301, 2003. 2, 3
- [19] Yashoda Sharma, Lovisa Cheung, Kara K Patterson, and Andrea Iaboni. Factors influencing the clinical adoption of quantitative gait analysis technology with a focus on clinical efficacy and clinician perspectives: A scoping review. *Gait & Posture*, 108:228–242, 2024. 1
- [20] D Sutherland. The development of mature gait. *Gait & posture*, 6(2):163–170, 1997. 1
- [21] Tian Tan, Tom Van Wouwe, Keenon F Werling, C Karen Liu, Scott L Delp, Jennifer L Hicks, and Akshay S Chaudhari. Gaitdynamics: A generative foundation model for analyzing human walking and running. *Nature Biomedical Engineering*, pages 1–13, 2026. 1
- [22] Johanna Vielemeyer, Lisa Tronicke, Lucas Schreff, Rainer Abel, Knut Lechler, and Roy Müller. A full-body motion capture gait dataset of healthy young adults walking ramps up and down. *Scientific Data*, 2026. 1
- [23] Guojun Yun, Meihuan Huang, Jianguo Cao, and Xianming Hu. Selective motor control correlates with gross motor ability, functional balance and gait performance in ambulant children with bilateral spastic cerebral palsy. *Gait & Posture*, 99:9–13, 2023. 1, 2