Evaluating the Feasibility of Deploying Quantized Human Pose Estimation on Smartphone for Gait Analysis for-children with Cerebral Palsy

Rahul Islam[1], Seth Donahue[2], Jeremy Bauer[2], Susan Sienko[2], Anita Bagley[2], Joseph Krzak[2], Maura Eveld[2], Karen Kruger[2], Ross Chafetz[2], Vedant Kulkarni[2], Hyeokhyen Kwon[1,3]

[1]Emory University, Atlanta, GA, USA, [2] Shriners Children's, USA, [3] Georgia Institute of Technology, Atlanta, GA, USA

E-mail: hyeokhyen.kwon@emory.edu Web: https://kwonvitallab.github.io

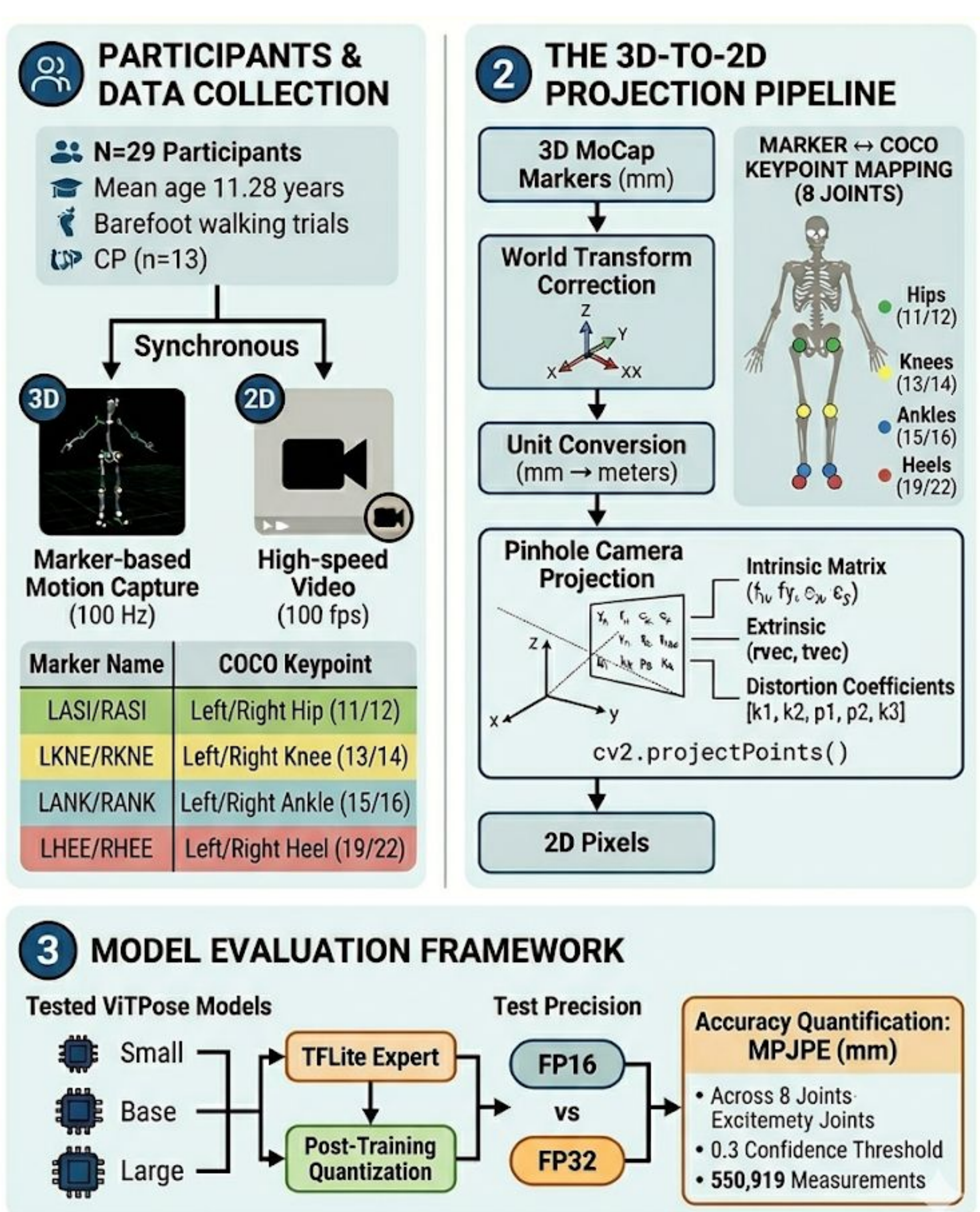
INTRODUCTION

Marker-based motion capture is the gold standard for clinical gait analysis but is expensive and limited to specialized gait labs, reducing accessibility for children with cerebral palsy (CP). Markerless motion capture with deep learning pose estimation could enable gait assessment using mobile devices in real-world settings [2,5]. ViTPose provides state-of-the-art accuracy for human pose esitmation [1], but mobile deployment often requires a pose estimation model that not only reduces computational costs but also minimizes quantization error with minimal reduction in precision [3]. The purpose of the current study is to evaluate whether 16-bit half-precision floating-point (FP16) ViTPose maintains clinically acceptableaccuracy relative to 32bit single precision floating point (FP32) in mobile devices.

CLINICAL SIGNIFICANCE

Mobile markerless gait analysis could improve access to gait assessment and enable longitudinal monitoring for children with CP outside clinical laboratories [3,4]. However, quantized models (a compressed version of an AI model) must be validated to ensure efficiency gains do not reduce accuracy enough to affect clinical decision-making [6].

METHOD



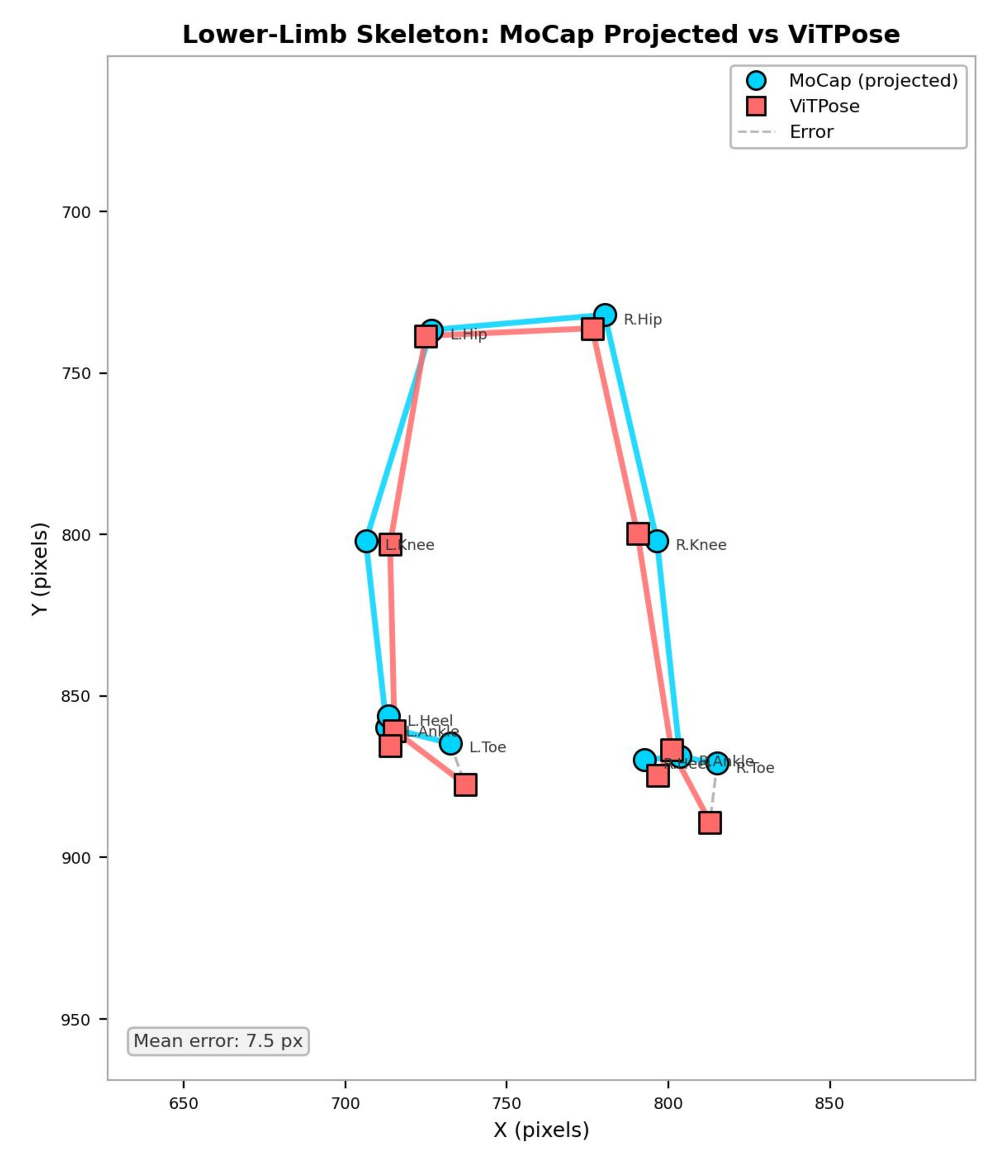


Figure 1. MoCap Projected vs ViTPose

RESULTS

Model	FP16 MPJPE (mm)	FP32 MPJPE (mm)	Mean Diff(mm)	% Diff
Small	40.14 ± 14.27	40.14 ± 14.27	-0.000051	-0.00013%
Base	39.77 ± 14.35	39.77 ± 14.35	+0.000034	+0.00009%
Large	39.76 ± 14.43	39.76 ± 14.43	+0.002076	+0.00522%

Table 1. The FP16-FP32 precision difference was negligible across all model sizes and joints

Joint	MPJPE (px)	N Measurements
Heels	30.5–30.6	44,662
Ankles	33.7–34.9	44,662
Knees	44.9–45.6	44,662
Hips	48.7–49.2	44,608

Table 2. Joint-wise accuracy (Large model FP16)

DISCUSSION

FP16 quantization caused negligible accuracy loss (<0.003 mm) compared to FP32 in ViTPose-based lower-extremity pose estimation, indicating suitability for mobile deployment. The quantization error was ~6 orders of magnitude smaller than the overall MPJPE (39.76–40.14 px; ~20–25 mm), consistent with prior markerless gait studies [2,5]. ViTPose Small achieved comparable accuracy to larger models with lower computational cost, making it ideal for edge deployment. While evaluated in controlled 2D lab conditions, future work should assess real-world performance.

REFERENCE

[1] Xu Y, Zhang J, Zhang Q, Tao D. ViTPose++: Vision Transformer for Generic Body Pose Estimation. IEEE Trans Pattern Anal Mach Intell. 2024;46(1):475-489.

[2] Cheng M, et al. Accuracy, Validity, and Reliability of Markerless Camera-Based 3D Motion Capture Systems versus Marker-Based 3D Motion Capture Systems in Gait Analysis: A Systematic Review and Meta-Analysis. Sensors. 2024;24(11):3686.

[3] Wishaupt T, et al. The applicability of markerless motion capture for clinical gait analysis in children with cerebral palsy. Sci Rep. 2024.

[[4] Anderson JT, Stenum J, Roemmich RT, Wilson RB. Validation of markerless video-based gait analysis using pose estimation in toddlers with and without neurodevelopmental disorders. Front Digit Health. 2025.

[5] D'Haene W, et al. Validation of a 3D Markerless Motion Capture Tool Using Multiple Pose and Depth Estimations for Quantitative Gait Analysis. Sensors. 2024;24(22):7105.

[6] Ali MM, Hassan MM, Zaki M. Human Pose Estimation for Clinical Analysis of Gait Pathologies. Evol Bioinform Online. 2024;20:11779322241231108.