Kinematic Estimation Using Deep Learning and IMU Data

Brough, Chi-Yeh Chen (Derek), Kwon, Mahajan, Tang

1. Abstract

Accurate estimation of lower-limb joint kinematics is crucial for diagnosing gait abnormalities, rehabilitation, and sports biomechanics. Traditional laboratory-based motion capture systems, while precise, are expensive and limited to controlled settings. Recent advances in inertial measurement units (IMUs) combined with deep learning models, have offered promising, cost-effective alternatives. However, these methods struggle with limited generalizability across diverse datasets due to differences in sensor placement, orientation, and inherent noise characteristics, highlighting a significant gap in achieving robust, real-world applicability. To bridge this gap, we developed and tested a transformer-based deep learning model enhanced by Rotary Positional Embedding (RoPE), hypothesizing that transformers would outperform existing convolutional architectures devised by Rapp Et Al through superior temporal and spatial modeling capabilities; this hypothesis consequently defined our success criteria. We harmonized synthetic IMU data (CMU-AMASS dataset), real IMU data (CMU lab dataset), and baseline data (Calgary dataset) via multiple preprocessing steps including resampling, orientation alignment, placement alignment, time alignment, and noise equalization. Although our transformer model alone did not surpass the CNN baseline in accuracy, fine-tuning with harmonized datasets significantly improved generalization performance, reducing prediction errors across multiple joint movements. These findings underscore that model complexity alone is insufficient for accuracy enhancements; instead, integrating diverse datasets substantially moves the needle towards creating robust, deployable, IMU-based kinematic estimation systems suitable for clinical and real-world scenarios.

2. Background & Introduction

Understanding lower limb joint kinematics is crucial for assessing human movement, diagnosing gait abnormalities, and monitoring rehabilitation progress. Research on lower limb joint kinematics using inertial measurement units (IMUs) has advanced, offering a viable alternative to traditional laboratory-based analysis systems and making gait analysis more accessible for clinical applications [1]. Despite these advancements, challenges remain, including sensor drift, magnetic disturbances, and the need for precise sensor-to-segment alignment[2]. To combat such issues, various computational strategies have been developed.

Traditional strategies rely on sensor fusion and biomechanical models. McGrath et al. [5] developed a method using principal component analysis (PCA) of thigh-shank angular velocities during natural gait initiation to estimate knee flexion-extension axes. This approach achieves approximately 3.49° zero-mean RMSE without manual alignment. In sensor fusion, Potter et al. [6] demonstrated an error-state Kalman filter (ErKF) that maintains hip angle errors below 1.4° during walking. Their method uses stance-phase zero-velocity updates (ZUPTs) and joint range-of-motion constraints to reduce drift in multi-IMU arrays, although magnetic disturbance mitigation is achieved by excluding magnetometers entirely.

More recently, machine learning (ML) methods have gained traction. Machine learning methods reproduce the results of traditional methods, but require large datasets, with joint angle prediction error decreasing with increasing dataset size. [2]

Rapp et al. [2] presented a magnetometer-independent framework using synthetic IMU data (accelerometer and gyroscope only) to train hybrid CNN-LSTM models. Their method achieves 1.27°–3.34° RMSE across planes during walking and running, outperforming traditional methods in unconstrained environments. Another significant contribution by Mundt et al. [3] involves comparing CNN, MLP, and LSTM networks for estimating lower-limb joint angles and moments. They found CNNs excelled in joint angle estimation (27-35% lower normalized RMSE than MLPs) by leveraging spatial correlations between adjacent IMUs through 1D convolutions.

Building upon the framework developed by Rapp et al. [2], transformers present a compelling solution to address persistent challenges in IMU-based kinematic estimation. Transformers' parallelizable self-attention mechanisms with the ability to attend to global context may overcome temporal modeling limitations inherent in LSTM architectures used by Mundt et al[3] while maintaining computational efficiency required for real-time applications [9,10]. Additionally, transformers have multiple attention heads that can learn different aspects of the signal concurrently.

A critical gap in current IMU-based models is generalizability across different subjects and conditions. Many models are trained and tested on a single dataset (often with fairly homogeneous demographics or gait tasks) – for instance, Mundt et al.[3] validated via leave-one-subject-out on their dataset, and Rapp et al.[2] trained on a synthesized dataset derived from one motion-capture repository. This can limit performance on truly independent populations. Rapp's method has yet to be tested on real IMU data and different cohorts.

Thus, the *open research question* we aim to address is three-fold:

1. Can a transformer-based deep learning framework estimate lower-extremity joint angles from IMU data with performance exceeding the baseline established by CNN (Rapp et al. [2])?
2. How can we harmonize diverse datasets with varying sensor placement, subjects, and datatype?
3. Can the transformer model demonstrate robustness to sensor variations and inter-subject differences across diverse datasets?

3. Methodology

3.1 Reproducing Baseline CNN & Proposed Transformer Model

3.1.1 CNN

We reproduced the Conv1D baseline from Rapp et al. (2021) as a joint‑specific prediction task: for each joint (hip, knee, ankle), the model takes a 16-dimensional input at each time step (two adjacent segments × eight IMU channels: acc_x, acc_y, acc_z, |acc|, gyro_x, gyro_y, gyro_z, |gyro|) and outputs three kinematic angles (flexion/extension, abduction/adduction, internal/external rotation). Using the 1D convolution neural network architecture, we stacked 1D‑convolutional layers with padding to preserve sequence length, followed by fully connected layers; we then optimized learning rate, dropout, layer sizes, and kernel widths via 300 TPE trials in hyperopt (_11_optimize_hyperparams.py), training in two phases (short windows for local features, longer windows for full gait cycles) with MSE loss, Adam optimizer, and exponential LR decay. Our implementation achieved RMSEs on Calgary walking and running data matching the published baseline, confirming a successful reproduction.

3.1.2 Transformer:

In the exact same settings as described above, we employed a lightweight transformer with rotary position embeddings(RoPE)(pure_transformer.py) that ingests a 16-dimensional input at each time step (two adjacent segments × eight IMU channels and outputs three kinematic angles per time step. The core of the Transformer is a stack of N=4 encoder layers. Each transformer block consists of (1) a multi-head self-attention with H=6 heads, using RoPE to incorporate position, followed by (2) a position-wise feedforward network. The attention output and the feedforward outputs both use residual connections followed by layer normalization. The feedforward block expands the dimension to 256 before projecting back to the original dimension, with intermediate ReLU and dropout (dropout probability 0.1). Finally, after stacking 4 such layers, we apply an output projection at each time step to yield the predicted 3 kinematic angles.

3.2 Datasets and Data Harmonization

In our project, we merged multiple datasets — synthetic IMU Calgary Running Clinic dataset used by Rapp et al.[2], the CMU AMASS motion capture dataset [11], and real IMU data collated at Eni Halilaj 's research lab. Dataset information is included below *(Table 1)*.

Table 1: Type, activity, and subject count for merged datasets used during training and finetuning.

Dataset	Type	Activities	Subject Count
Calgary	Synthetic	Running	1015
		Walking	883
CMU AMASS	Synthetic	Running	13
		Walking	53
Halilaj	Real	Running	1
		Walking	3

To prepare the datasets for training and evaluation, we implemented a comprehensive data harmonization pipeline designed to address discrepancies in sampling rate, sensor placement, and orientation. Harmonizing data from different sources was necessary to ensure that the model could generalize effectively across datasets.

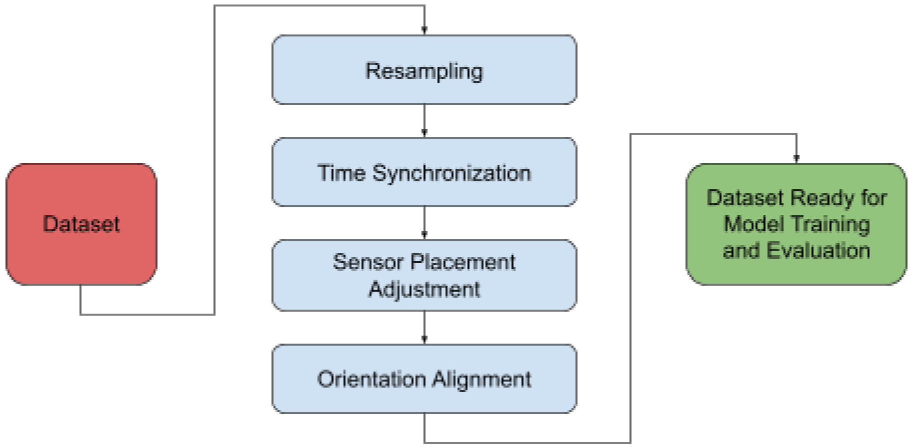

Figure 1: Data harmonization pipeline for diverse IMU datasets.

- *Resampling to Match Sampling Rates*

  We standardized all IMU and MoCap data streams to a common sampling rate by computing a sampling factor for each sequence (target frequency ÷ original frequency) and performing upsampling via SciPy's `interp1d`. Specifically, we reconstructed each signal's original time axis from its timestamps, generated a new uniform time grid at the target frequency, and applied `interp1d` with linear interpolation to compute resampled values. This approach preserved critical gait features, such as heel strikes and joint transitions, ensuring consistency across datasets for direct integration and comparison.

- *Time Synchronization*

  To align IMU and MoCap signals recorded on separate clocks, we implemented a peak-based shifting algorithm. First, we detected the three distinct peaks corresponding to calibration hops in both IMU and MoCap signals. We then calculated the temporal offset by comparing the index of the first IMU peak to the corresponding MoCap peak. Applying this offset to the IMU timeline synchronizes the two data streams. We validated the alignment by checking the remaining hop peaks for any residual drift, which was found to be minimal. Finally, we cropped out the calibration hops to isolate the synchronized activity segments for further analysis.

- *Orientation Alignment and Axis Transformation*

  Ideally, we would apply coordinate transformations based on each dataset's documented IMU information (sensor axes definitions and mounting orientations). However, many datasets lacked complete metadata, making this approach infeasible. Instead, we developed an experimental alignment method: for each dataset and sensor location, we plotted representative IMU signals (e.g., accelerometer and gyroscope channels) alongside the reference dataset's signals. By visually inspecting phase and sign differences, we determined whether axes needed swapping or sign flipping. From these observations, we constructed a transformation matrix for each sensor configuration. Once defined, these matrices were applied uniformly across all subjects in that dataset, ensuring consistent orientation alignment for downstream model training.

3.4 Fine-tuning

After performing data harmonization, we evaluated our pre-trained transformer model on the external lab dataset to test its generalizability. As shown in *Figure 2*, the model exhibited poor predictive performance on the harmonized external datasets, particularly on real-world lab data. To address this issue, we adopted a fine-tuning strategy aimed at improving cross-domain generalization. Specifically, we applied full-parameter fine-tuning—where all model weights were retrained without freezing any layers—starting from a model originally trained on the Calgary dataset.

We conducted fine-tuning on two external datasets: the AMASS dataset (synthetic motion capture data) and a real-world lab dataset provided by Eni's lab. The lab dataset consisted of 2.5-minute walking trials from two subjects. To increase the diversity of training samples, each trial was segmented into four parts, yielding six fine-tuning cases and two held-out test segments. For the AMASS dataset, 47 samples were used for fine-tuning and three for testing.

Our analysis focused on predicting three primary knee joint movements during walking: flexion/extension, adduction/abduction, and internal/external rotation.

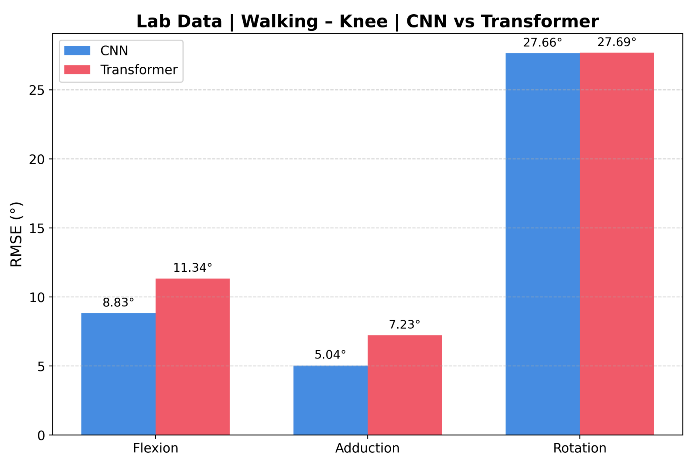

Figure 2: RMSE Comparison of Pre-trained CNN and Transformer Models on External Lab Walking Data (Knee Joint Angles)

3.5 Code

Our code used for this project can be found in [this] GitHub repository. We hope that by making our code available, future researchers can expand and improve upon our work. Specifically, our contributions towards the Rapp Et Al repository are as follows:

- Implementation of a custom Transformer model in nn_models/models/pure_transformer.py.
- Implementation of a custom fine-tuned process in data preprocessing/finetune.ipynb

4. Results

4.1 Baseline CNN & Transformer with RoPE Comparison

Our initial exploration compared a transformer-based deep learning model to a baseline convolutional neural network (CNN) using the Calgary dataset. We observed that, while transformer predictions closely mirrored those of the baseline Conv1D CNN model, the transformer model did not outperform the CNN.

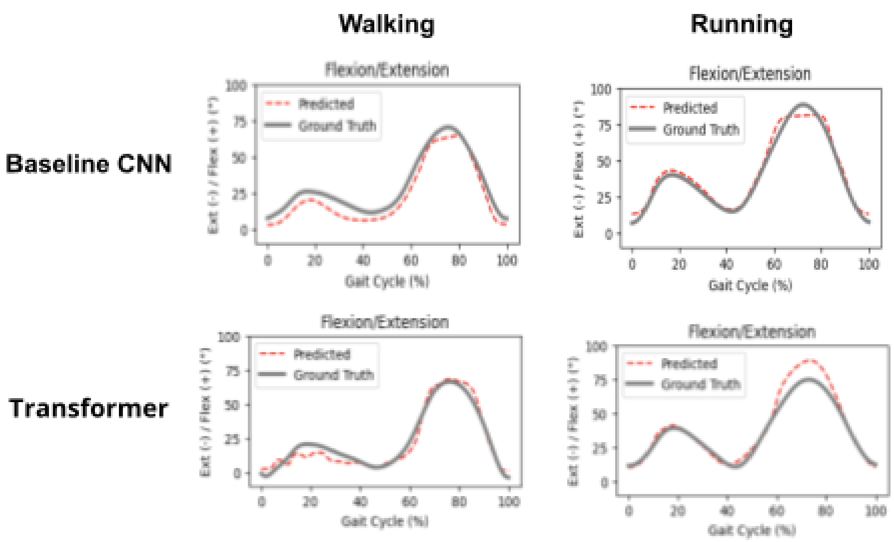

Figure 3: Knee angle prediction for baseline CNN and transformer model trained and tested on the Calgary dataset.

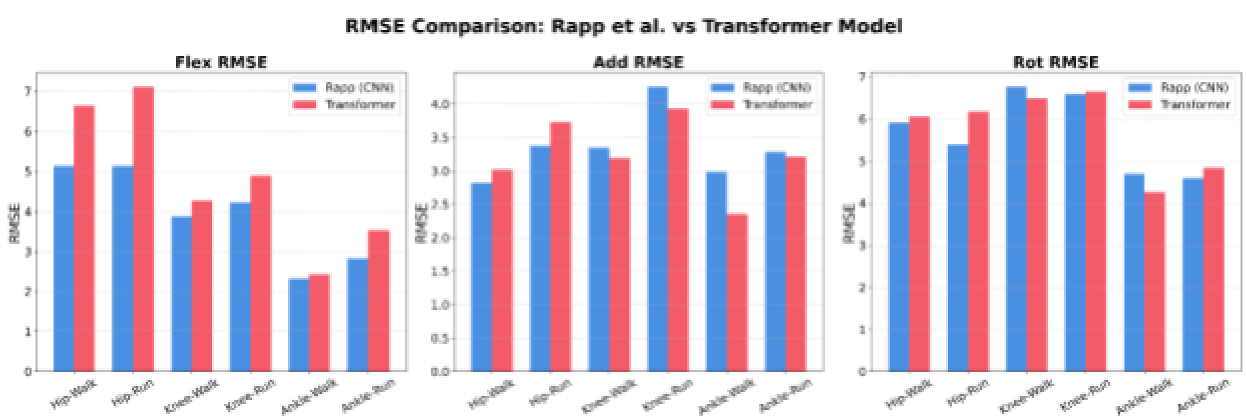

Figure 4: A comparison of RMSE values for baseline and transformer models trained and tested on the Calgary dataset.

While the transformer seemed to outperform the CNN in adduction, it underperformed in flexion, and had mixed results for rotation. These discrepancies, or rather lack of discrepancies, exhibit the limitations of the transformer in generalizing to unseen real IMU data, indicating that merely increasing the complexity of the model architecture was insufficient for enhancing performance. We thus turned to our data harmonization efforts in hopes that increasing dataset size and diversity would improve model performance.

4.2 Data Harmonization Results

To address this generalization problem and bridge the diverse dataset gap, we implemented extensive harmonization steps (as outlined in our methods) on synthetic data from CMU AMASS and real data from the Halilaj dataset. Harmonizing these datasets proved essential due to discrepancies in sensor orientation and placement, which were successfully managed using a combination of manual and computational techniques. *(Figure 5)*

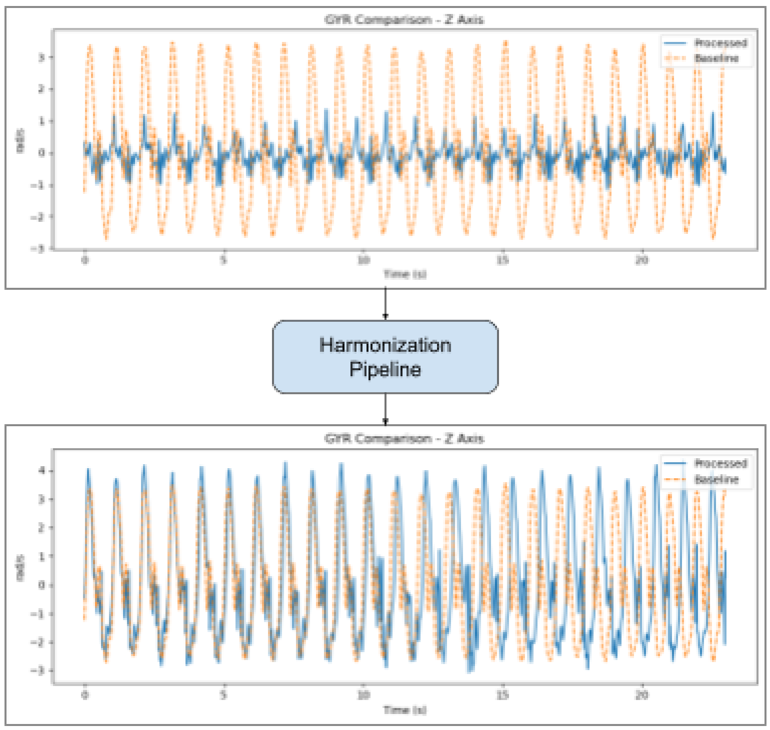

Figure 5: Halilaj subject data undergoing preprocessing steps. Blue indicates subject data and orange indicates a sample of gyroscope data from the Calgary dataset used for harmonization validation.

From the data harmonization process, we created a cheatsheet we hope will act as a beginner's guide for other researchers in need of IMU data harmonization guidance *(Table 2)*.

Table 2: Cheat Sheet for Harmonizing IMU / MoCap Data Sets.

	What to do	Why it matters	Quick check
Tip 1 Standardize measurement units.	Bring all linear acc. to m/s^2, gyro. to degrees/second.	Unit mismatches masquerade as drift or bias.	Verify that values should sit in plausible human-motion ranges (±20 m/s^2, ±500 deg/s).
Tip 2 Align coordinate systems.	Reorient sensors to an agreed axis convention (e.g., Z-up, X-forward).	Misaligned axes produce phase-shifted or inverted signals, preventing dataset integration.	Overlay a representative gait cycle from each set; correctly aligned signals coincide in phase and morphology.
Tip 3 Match sampling rates.	Resample all streams to a common frequency.	Keeps stride features intact, avoids aliasing, and simplifies multiset comparison.	Frequency domains (e.g., FFT) on resampled signals show gait peaks at the same bin for all sets.
Tip 4 Record sensor placement.	Check the alignment of virtual/real IMU sensor placement.	Reproducibility depends on consistent attachment locations across datasets.	Store schematic diagrams showing each sensor's local axes and anatomical landmarks.
Tip 5 Induce intentional noise.	Add Gaussian & bias drift to synthetic sets, denoise real sets with the same filter.	Equalises noise across sets(real/synthetic) so the model does not over-trust a single source.	Train-time SNR ~ test-time SNR (estimate via power-spectral density).
Tip 6 Suppress differentiation spikes (synthetic).	Remove and apply the Savitzky-Golay filter or the light Kalman filter.	Differentiation amplifies high-frequency noise, creating edge artifacts and spurious peaks.	Identify second derivative values exceeding ±6 σ.
Tip 7 Keep folder structure consistent.	Keep a strict folder structure across datasets.	Consistent file management simplifies feeding data to the model.	A folder structure could include TASK/Subject/Joint for the joint angles and TASK/Subject/IMU_Location.

4.3 Fine-tuning Results

With our harmonized dataset, we implemented fine-tuning on the transformer model using the AMASS dataset. This fine-tuning approach improved both RMSE and robustness, enhancing the model's adaptability to external datasets. With just one pre-trained model, we can fine-tune it on a portion of any new harmonized dataset to enable more accurate joint movement predictions tailored to that specific data source.

We noted substantial improvements in predicting knee joint movements during walking *(Figure 6)*.

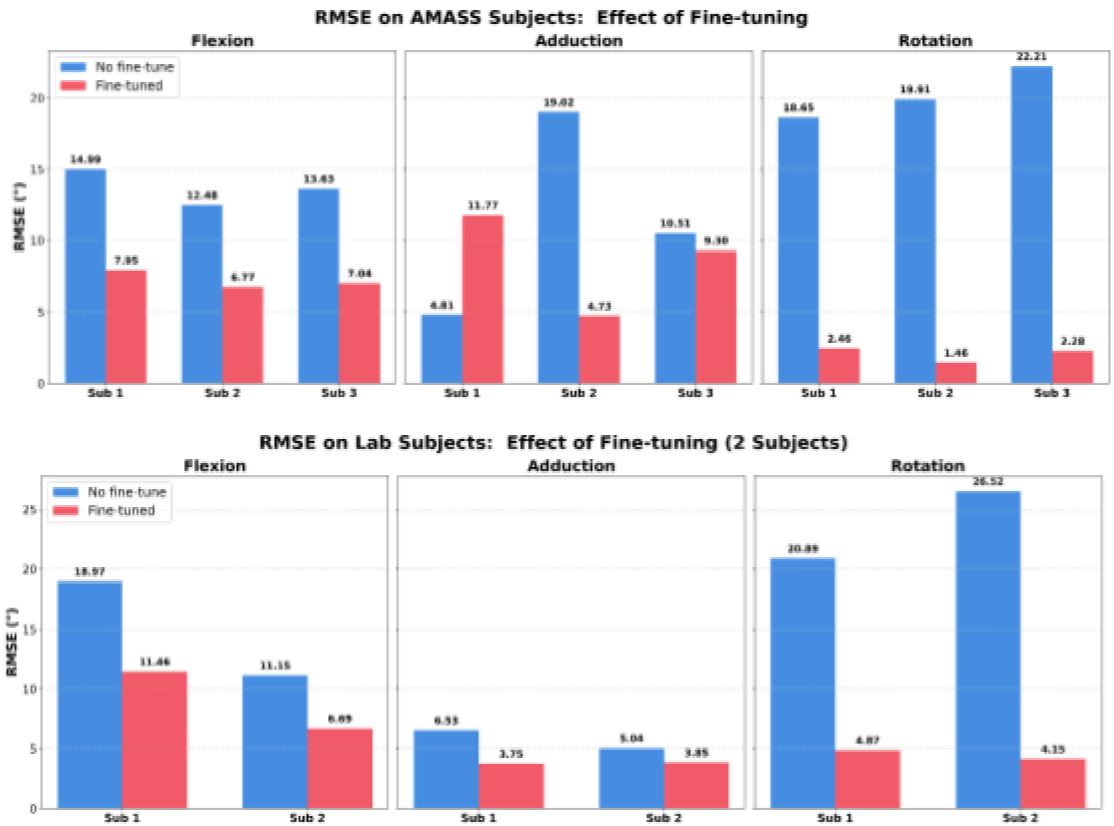

Figure 6: Comparison of RMSE values before and after fine-tuning on external datasets. Blue bars indicate performance without fine-tuning, while red bars indicate performance after fine-tuning. Top: AMASS dataset (synthetic data); Bottom: Lab dataset (real-world data). Results are shown for three knee joint movements: flexion/extension, adduction/abduction, and rotation.

Fine-tuning on AMASS and Halilaj datasets reduced the model's error significantly, demonstrating that targeted retraining on specific data sources can notably enhance model performance. The most significant reduction in error occurred for rotation with error being reduced by an average of 89.7% across subjects for fine-tuning on the AMASS dataset and 80.9% across subjects for fine-tuning on the Halilaj dataset.

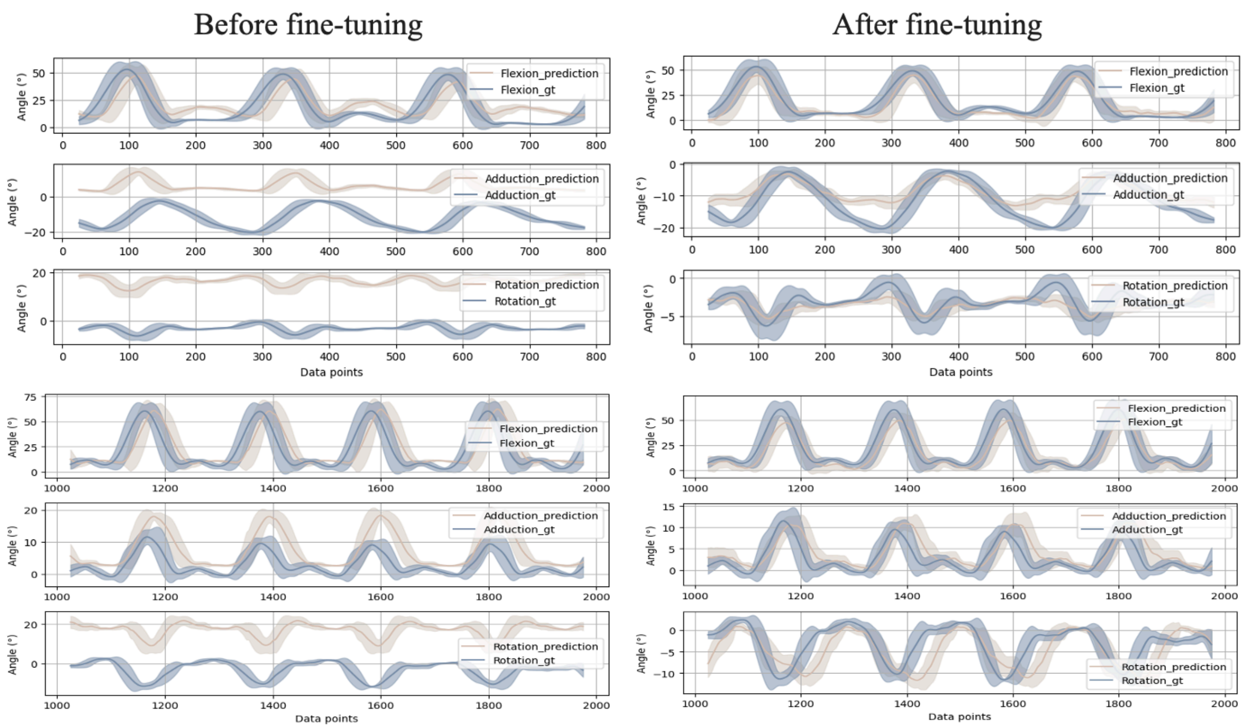

Figure 7: Predicted versus ground truth joint angles before and after fine-tuning on external datasets. Each row shows prediction results for a specific joint movement: flexion/extension, adduction/abduction, and rotation. Blue lines indicate ground truth (gt), and orange lines show model predictions. Top: gait cycles of AMASS data. Bottom: gait cycles of Halilaj data.

Qualitatively, this improved prediction accuracy is shown well on the gait plots *(Figure 7)*, with model predictions closely matching ground truth.

However, fine-tuning also revealed potential limitations. The accuracy of the transformer model still suffered due to factors such as inconsistencies in sensor orientation, disparities between synthetic and real data, IMU types, placements, and limited external data availability. Though our results are promising, much work is to be done to simplify and ensure accuracy in the data harmonization process.

5. Discussion - Addressing the Research Questions

Our study set out to investigate whether increased model complexity alone could improve IMU-based kinematic estimation, whether diverse datasets could be integrated effectively, and if more data indeed contributes to better performance.

5.1 Model Complexity vs Data

Our findings indicate that increased model complexity alone does not guarantee improved performance. Despite the sophisticated self-attention mechanisms in the transformer model, it initially did not surpass the simpler baseline CNN model in accuracy. Hence, *we couldn't meet our success criteria of* *outperforming the baseline CNN* [2].This highlights a crucial insight — model complexity must be appropriately matched by sufficient, high-quality, and diverse data to realize tangible performance gains. However, exploring various architectural modifications to our transformer model through extended experimentation might provide further insights into optimizing performance.

Additionally, while our findings contradict the learnings from *Team-2,* it is important to note the nature of our research is very different from theirs. While we are using a transformer model for prediction of joint angles, their task was one of classification. Apart from the difference in task, the entirety of the project pipeline ranging from the datasets used to the model architectures implemented is very different. In order to have a reliable comparison, setting the initial pipeline similarly for both of the teams is critical. However, *Team-3* finds the key learnings from *Team-2* to be quite interesting. The use of an autoencoder network to learn a meaningful smaller representation of the input for activity classification is impressive. While we could have utilized a similar network and proceeded to work along the same lines, we believe that encoding the input for a regression task is suboptimal and risks oversimplifying critical information necessary for precise angle predictions in a regression task setting. However, the mutual consensus among our team also suggests that testing the same autoencoder model and adapting it for a regression task presents an intriguing learning opportunity.

5.2 Data Integration and Harmonization

Integration and harmonization of diverse datasets emerged as pivotal elements in enhancing our model's performance. Successfully merging synthetic (Calgary and AMASS CMU) and real-world (Halilaj Lab) IMU datasets posed significant technical challenges due to inherent differences in sampling rates, sensor placements, orientations, and data noise characteristics. We addressed these barriers through a systematic preprocessing pipeline encompassing resampling, temporal synchronization, orientation alignment, and noise equalization. Our results indicate that carefully harmonized data significantly reduce discrepancies during model evaluation and enhance the generalization capacity across datasets.

We encountered notable practical difficulties in data harmonization due to missing metadata, incomplete documentation, and varying sensor hardware across datasets. To overcome these, we developed pragmatic solutions, including empirical alignment techniques involving visual inspection and iterative experimentation. These methods were effective but required extensive manual intervention, highlighting opportunities for streamlining. To improve future applicability, we propose developing a universally applicable automated harmonization framework capable of dynamically identifying optimal transformations based on minimal metadata and data-driven heuristics.

5.3 Moving Towards Generalizability & Robustness

A strength of our study lies in demonstrating how strategic fine-tuning on harmonized datasets substantially enhances model robustness and generalization performance. Fine-tuning the transformer model on external datasets resulted in significant reductions in prediction error across joint movements.

Our experiments emphasize that generalization does not inherently follow from increased model complexity alone but rather from training on sufficiently diverse data reflective of real-world variations. To foster broader adoption of IMU-based systems in clinical practice, the community must prioritize creating and sharing harmonized, large-scale, multi-source datasets. Clearly defined benchmarks for evaluating generalization across diverse external datasets would further facilitate consistent and meaningful comparisons between studies. By establishing a standard, researchers can reliably quantify how well models generalize beyond their initial training environments.

5.4 Limitations and Future Work

While our work provides valuable insights into model-data relationships and dataset harmonization, it also has several limitations. Foremost, our external datasets used for fine-tuning and validation were limited in both size and variability. The real-world dataset, notably from the CMU Lab, included data from very few subjects, limiting our ability to generalize broadly. Consequently, despite notable improvements after fine-tuning, our results might still not fully reflect performance across a broader population or diverse movement scenarios.

Furthermore, discrepancies in sensor placements caused inaccuracies in angle predictions, despite the harmonization attempts. Sensor hardware differences (e.g: IMU(synthethic/real), sampling rates, and measurement noise levels) between synthetic and real datasets remain challenging to address completely, suggesting room for refinement in our harmonization strategy.

To address these limitations, we propose several targeted avenues for future work:

- *Automated Harmonization Framework*: Developing automated data-harmonization scripts leveraging machine learning or heuristic optimization could minimize manual alignment and standardize preprocessing. Such frameworks should handle varying metadata completeness and identify the necessary transformations automatically once the dataset is fed to the framework.
- *Comprehensive Benchmarking*: Establishing benchmarks and testing frameworks across diverse datasets would significantly clarify model robustness. Large-scale benchmarking could help identify conditions under which models succeed or fail, providing clearer guidelines for clinical applications.
- *Few-shot and Transfer Learning Approaches*: Investigating few-shot and transfer learning methods, where minimal subject-specific data enables rapid model adaptation, could dramatically enhance model usability in practical scenarios. Such methods would enable clinicians to quickly fine-tune generalized foundational models on limited patient-specific data, improving individualized prediction accuracy.
- *Dataset Size and Quality Thresholds*: An essential unresolved question pertains to quantifying how much and what quality of external data are necessary to ensure robust model generalization. Systematic studies addressing this would provide concrete guidelines for future data collection.

Collectively, addressing these limitations through focused research initiatives will significantly advance the field, moving closer toward practical and clinically viable IMU-based kinematic estimation solutions.

5.5 Overall Impact

In conclusion, our study provides meaningful incremental contributions to wearable health technology research, emphasizing practical methodologies for data harmonization and reinforcing the necessity of data adequacy alongside model sophistication.

6. Acknowledgements

We sincerely thank our Teaching Assistant, Zhixiong (Jack) Li, for his invaluable support, guidance, and technical insights throughout this project. We also extend our heartfelt gratitude to Prof. Eni Halilaj for her mentorship, thoughtful feedback, and encouragement that significantly shaped our research direction. Their combined efforts greatly enhanced our learning experience and contributed profoundly to our project's successful completion.

References

[1] Picerno, Pietro. “25 Years of Lower Limb Joint Kinematics by Using Inertial and Magnetic Sensors: A Review of Methodological Approaches.” Gait & Posture 51 (January 1, 2017): 239–46.

[2] Rapp, Eric, Soyong Shin, Wolf Thomsen, Reed Ferber, and Eni Halilaj. "Estimation of kinematics from inertial measurement units using a combined deep learning and optimization framework." Journal of Biomechanics 116 (2021): 110229.

[3] Mundt, Marion, William R. Johnson, Wolfgang Potthast, Bernd Markert, Ajmal Mian, and Jacqueline Alderson. "A comparison of three neural network approaches for estimating joint angles and moments from inertial measurement units." Sensors 21, no. 13 (2021): 4535.

[4] Xiang, Liangliang, Alan Wang, Yaodong Gu, Liang Zhao, Vickie Shim, and Justin Fernandez. "Recent machine learning progress in lower limb running biomechanics with wearable technology: A systematic review." Frontiers in Neurorobotics 16 (2022): 913052.

[5] McGrath, T.; Fineman, R.; Stirling, L. "An Auto-Calibrating Knee Flexion-Extension Axis Estimator Using Principal Component Analysis with Inertial Sensors." Sensors 18 (2018): 1882.

[6] Potter, M.V.; Cain, S.M.; Ojeda, L.V.; Gurchiek, R.D.; McGinnis, R.S.; Perkins, N.C. "Evaluation of Error-State Kalman Filter for Lower-Limb Kinematic Estimation." PLOS ONE 16(4) (2021): e0249577.

[7] Al Borno, Mazen, et al. "OpenSense: An open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations." Journal of Neuroengineering and Rehabilitation 19, no. 1 (2022): 22.

[8] Hur, Benjamin, Sunin Baek, Inseung Kang, and Daekyum Kim. “Learning Based Lower Limb Joint Kinematic Estimation Using Open Source IMU Data.” Scientific Reports 15, no. 1 (2025): 5287.

[9] Ioffe, Sergey, and Christian Szegedy. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” ICML (2015): 448–456.

[10] Feichtenhofer, Christoph, Haoqi Fan, Jitendra Malik, and Kaiming He. “SlowFast Networks for Video Recognition.” IEEE ICCV (2019): 6201–6210.

[11] The CMU AMASS motion capture data used in this project was obtained from mocap.cs.cmu.edu.

Table A: Contributions of Group Members To Final Report

Section	Contributor(s)	Notes
1. Abstract	Abhinit, Derek	-
2. Background & Introduction	Abhinit, Elle	Large majority taken from Milestone 1; major background research conducted by Elle
3.1.1 CNN	Abhinit	-
3.1.2 Transformer	Abhinit	Initial idea proposed by Abhinit
3.2 Datasets and Data Harmonization	Jonathan, Elle	-
3.4 Fine-tuning	Derek	-
3.5 Code	Abhinit, Owen, Elle	Github organization by Owen, Derek
4.1 Baseline CNN & Transformer with RoPE Comparison	Jonathan, Elle, Owen, Derek	Major Implementation and figure by Owen
4.2 Data Harmonization Results	Jonathan, Elle, Derek	Data Processing (Lab data) by Derek and Jonathan; AMASS processing by Abhinit and Elle; Original table and figure organization by Owen, Jonathan and Derek
4.3 Fine-tuning Results	Derek, Jonathan	Major fine-tuning Implementation by Derek; Figure created by Derek and Owen
5.1 Model Complexity vs Data	Abhinit, Elle	-
5.2 Data Integration and Harmonization	Abhinit	-
5.3 Importance of Data Quantity and Quality	Abhinit	-
5.4 Limitations and Future Work	Abhinit	-
5.5 Overall Impact	Abhinit	-