Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 756–765
https://doi.org/10.54941/ahfe1008308

EdgeConv-Based Spatial Difference Extraction of Plantar Pressure Improves Mediolateral Ground Reaction Force Estimation During Walking

Hideyuki Nagashio[1], Yusuke Osawa[1], Kazunori Kaede[1,2], and Keiichi Watanuki[1,2]

[1]Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan
[2]Advanced Institute of Innovative Technology, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338-8570, Japan

ABSTRACT

Falls and fractures are major concerns in super-aged societies, highlighting the need for gait assessment in daily living environments. Wearable footwear devices equipped with plantar pressure sensors and inertial measurement units (IMUs) enable remote gait monitoring without laboratory-based equipment; however, they cannot directly measure three-dimensional ground reaction forces (3D GRFs), and estimating the mediolateral component (F_x), which is related to balance control, remains challenging. This paper proposes a deep learning model that integrates Edge Convolution (EdgeConv), a graph neural network layer, to extract spatial features from plantar pressure data and improve F_x estimation. Gait experiments were conducted with 12 healthy participants walking on an instrumented treadmill, during which 3D GRFs and wearable sensor data from eight plantar pressure sensors and a six-axis IMU were synchronously recorded. The proposed model combines EdgeConv for spatial feature extraction with a Transformer for temporal feature learning and was evaluated using subject-independent six-fold group cross-validation. For F_x estimation, the proposed model achieved an NRMSE of $8.79 \pm 1.26\%$ and an R^2 of 0.76 ± 0.09, compared with $9.30 \pm 1.52\%$ and 0.73 ± 0.09 for the Transformer and $9.47 \pm 1.77\%$ and 0.71 ± 0.13 for the GCN. These results indicate that learning local spatial features with EdgeConv enhances mediolateral GRF estimation from wearable footwear sensor data.

Keywords: 3D Ground Reaction Forces (3D GRFs), EdgeConv, Walking analysis, Plantar pressure, IMU sensor

INTRODUCTION

In Japan, the rising number of individuals requiring long-term care presents a societal challenge. Among older adults in the pre-care-dependency stage who require support, the leading contributing factors are age-related frailty (17.7%), joint disease (14.8%), and fractures or falls (14.7%), and factors associated with declining walking ability account for approximately half of all cases (Ministry of Health, Labour and Welfare, Government of Japan,

Received October 17, 2026; Revised November 22, 2026; Accepted December 2, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

2025). Therefore, preventing progression to care dependency requires early detection of declining walking ability and timely interventions.

The ground reaction force (GRF) directly reflects the mechanical loads acting on the body and serves as an important quantitative indicator of gait stability. In particular, the three-dimensional ground reaction force (3D GRF) provides useful information for evaluating propulsion and balance control during walking (Bowden et al., 2006; Bauby et al., 2000). However, GRF measurements using force plates are limited to laboratory settings and are therefore unsuitable for continuous gait monitoring in daily life. Moreover, the awareness of stepping on a force plate, along with the artificial nature of the laboratory environment, can affect gait speed and patterns, posing a challenge to evaluating natural walking in daily life (Shah et al., 2020; Xia et al., 2026).

To overcome these limitations, shoe-type devices that combine plantar pressure sensors with inertial measurement units (IMUs) have attracted attention in recent years. Such devices can simultaneously measure the plantar load distribution and foot kinematics; however, obtaining the 3D GRF directly from sensor data through physical transformation alone remains difficult.

Therefore, studies estimating GRF from sensor data acquired by shoe-type devices using deep learning have proliferated (Kim et al., 2023). These studies commonly report that the mediolateral component (F_x), which is critical for balance control and lateral sway, is estimated less accurately than the vertical component (F_z), which supports body weight, or the anteroposterior component (F_y). For example, a study that estimated GRF using a many-to-one regression model, which is a long short-term memory (LSTM) network trained separately for each axis, from stance-phase plantar pressure insole data reported a normalized root mean square error (NRMSE) of 5.50 ± 2.25% for the vertical component and 6.26 ± 2.52% for the anteroposterior component, compared with 12.96 ± 4.98% for the mediolateral component. Clearly, the estimation accuracy for the mediolateral component is markedly lower than that of the other two components (Hajizadeh et al., 2023).

A possible reason for the difficulty in estimating the mediolateral component is that the underlying mechanical information is closely related to the left–right asymmetry in the plantar pressure distribution and to local load transfer within the sole. Most conventional deep learning models use the pressure value from each sensor as an input feature without explicitly learning the spatial relationships between sensors or the pressure differences between adjacent sensors.

Based on the idea that the spatial distribution and gradient of the plantar pressure provide key information for estimating the mediolateral component, this paper proposes a deep learning model that learns the spatial characteristics of the plantar pressure using a graph neural network (GNN). Specifically, Edge Convolution (EdgeConv) is introduced as a spatial feature extraction layer to integrate the inherent features of each sensor with those of its neighboring sensors and extract local spatial features, after which a Transformer learns the temporal dependencies. The purpose of this study

was to examine whether learning the spatial difference features of the plantar pressure through EdgeConv improves the estimation accuracy of the mediolateral component compared with a conventional Transformer-based model and a standard graph convolutional network (GCN)-based model.

GRF ESTIMATION METHOD AND EXPERIMENTAL PROCEDURE

Participants and Ethical Considerations

This study was approved by the Research Ethics Committee for Human Subjects Research of Saitama University (approval no. R6-E-15), and informed consent was obtained from all the participants. Participants were 12 healthy male adults (age: 23.5 ± 1.2 years; mean shoe size: 26.9 ± 0.5 cm) whose habitual shoe size fell within the range of 26.0–28.0 cm, which matched the size of the shoe-type device (27.0 cm).

Experimental Protocol

To account for the variability in the walking speed observed in daily life, three speed conditions were evaluated relative to each participant's self-selected comfortable walking speeds: subjectively slow, subjectively comfortable, and subjectively fast. The participants first walked on a treadmill and adjusted its speed until it matched their habitual walking speed, which was defined as subjectively comfortable speed. The subjectively slow and fast speeds were set to 5% below and 5% above the comfortable speed, respectively. The three conditions were tested in the order of subjectively slow, subjectively comfortable, and subjectively fast. The mean walking speed across all conditions was 3.8 ± 1.0 km/h. Data were acquired simultaneously during treadmill walking. A force-plate-instrumented treadmill measured the 3D GRF (F_x, mediolateral component; F_y, anteroposterior component; F_z, vertical component) (Figure 1(a)). The shoe-type device equipped with plantar pressure sensors (eight channels per foot) and an IMU (six channels) (Figures 1(b) and (c)) measured the plantar load distribution and foot kinematic information. The plantar pressure sensors were positioned in the load-bearing regions of the heel, midfoot, and forefoot; Table 1 presents the sensor locations.

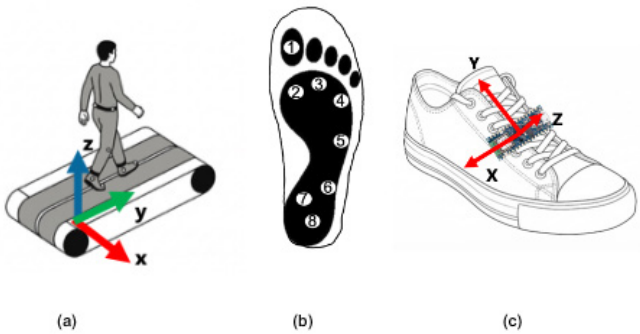


Figure 1: Overview of the experimental setup and measurement system: (a) Gait experiment on a treadmill; (b) Eight plantar pressure sensors installed on the foot; (c) Inertial measurement unit (IMU) placement.

Table 1: Plantar pressure sensor locations.

Node no.	Corresponding Location
1	Hallux
2	First metatarsal
3	Third metatarsal
4	Fourth metatarsal
5	Lateral arch
6	Lateral heel
7	Medial heel
8	Central heel

Data Preprocessing

The acquired data were preprocessed as follows. First, a fourth-order low-pass Butterworth filter was applied to remove electrical noise. The GRF was then normalized by body weight (%BW), plantar pressure data were normalized using min-max normalization, and IMU data were standardized using z-score normalization. The left foot data were mirrored to match the representation of the right foot, allowing a common model to be applied to the data from both feet. For segmentation, the stance phase of each gait cycle was extracted as a single stride, and each stride was time-normalized to 200 points before being inputted to the model. This process yielded a final dataset of approximately 13,000 strides.

GRF Estimation Models

The three models compared in this study differed progressively in how they incorporated spatial information. The Transformer-based model did not explicitly model the spatial arrangement of the sensors. The GCN-based model introduced adjacency relationships between sensors as a graph structure and incorporated spatial information by aggregating the features of neighboring sensors. The EdgeConv-based model explicitly learned the feature differences between sensors, captured local spatial variations in the plantar pressure, and examined their contribution to the estimation of the mediolateral component. Table 2 presents the detailed configurations of each model.

A GCN learns a feature representation that reflects the relationships between nodes on a graph composed of nodes and edges by aggregating the features of each node with those of its neighboring nodes. In this study, noting that the plantar pressure sensor arrangement forms a spatial structure fixed by anatomy, we represented each sensor as a node with a 1D node feature consisting solely of its pressure value and constructed a fixed graph in which the relationships between adjacent sensors served as edges. The coordinates of each sensor were not included in the node features; they were used only to compute the edge weights (the inverse Euclidean distance between adjacent sensors) such that information propagated more strongly between sensors that were close by, and the spatial features were extracted using the GCN.

EdgeConv was originally developed for the classification and segmentation of 3D point cloud data. A distinctive feature of EdgeConv is that in each layer, the graph is dynamically reconstructed via a k-nearest-neighbor search in the feature space, allowing the model to learn the local geometric structure within the point cloud. The output feature of node i under EdgeConv is generally defined as

$$x_{i'} = mean_\{j \in N(i)\} h_{\Theta}(x_i, x_j - x_i)$$

where $N(i)$ denotes the set of neighboring nodes of node i, and h_{Θ} is a nonlinear function with learnable parameters Θ (implemented in this study as a two-layer multilayer perceptron (MLP) comprising a linear layer, a rectified linear unit (ReLU) activation layer, and a second linear layer). The node features were limited to each sensor's pressure value (1D). Accordingly, the function takes as input the pressure value of the central node x_i, and its pressure difference from each neighboring node, $x_j - x_i$, and outputs a local edge feature. The edge features from the neighboring nodes are aggregated using the mean. Thus, rather than performing distance-based weighted aggregation as in the GCN, EdgeConv explicitly learns the pressure differences between adjacent sensors as an input.

Instead of dynamically reconstructing the graph in the feature space, as in the original formulation, we applied EdgeConv to the same fixed graph based on the anatomical structure of the sole used in the GCN-based model. That is, instead of learning a dynamic neighborhood structure, as in point-cloud processing, EdgeConv is applied within a predefined sensor-adjacency structure to learn the pressure difference between a central sensor and its neighboring sensors, which is a departure from its original application. This is intended to capture the local spatial variation in the plantar pressure, particularly information related to mediolateral load transfer.

Figure 2 shows the graph structures used in the GCN- and EdgeConv-based models. Each plantar pressure sensor is represented as a node, and the edges connect anatomically adjacent sensor locations. An additional edge spanning the medial longitudinal arch is defined to reflect the plantar load distribution more fully.

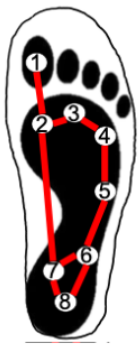


Figure 2: Graph structure of the plantar pressure sensor network.

Table 2: Model configuration.

Component	Transformer-Based Model	GCN-Based Model	EdgeConv-Based Model
Input representation	14-dimensional input (8-channel pressure + 6-channel IMU) projected to d_model = 128 via a linear layer	8-node graph from pressure (8 ch) separated from 6-channel IMU	
Spatial feature extraction	-	2-layer GCNConv (output dim. 16); edge weight = inverse Euclidean distance between nodes	2-layer EdgeConv (output dim. 16); edge feature [pressure value x_i, pressure difference $x_j - x_i$] transformed by an MLP; aggregation = mean
Feature fusion	-	8-node output (128 dim.) compressed to 32 dim. via a linear layer, concatenated with IMU features (6 dim.), then projected to $d_{\text{model}} = 128$ via a linear layer	
Positional encoding	sine/cosine functions		
Transformer encoder	3 layers, 4 heads, feedforward dim. 256, dropout rate 0.1		
Output layer	Linear(128, 64) → ReLU → Linear(64, 3)		

Training Conditions and Evaluation Metrics

The models were trained using the mean squared error (MSE) as the loss function, the Adam optimizer, a batch size of 512, and 200 training epochs. Model performance was evaluated using six-fold group cross-validation (CV) at the subject level. The 12 participants were divided into six folds, with each fold comprising eight participants for training, two for validation, and two for testing, allowing an assessment of the generalization performance of the model on unseen subjects. The evaluation metrics were the NRMSE and the coefficient of determination (R^2).

RESULTS

Tables 3 and 4 summarize the results of the between-subject evaluation obtained via the six-fold group CV. All three models—the Transformer-based, GCN-based, and EdgeConv-based models—achieved comparable estimation accuracies for the anteroposterior (F_y) and vertical (F_z) components.

In contrast, a different trend emerged for the mediolateral component (F_x). The Transformer-based model achieved an NRMSE of 9.30 ± 1.52% and an R^2 of 0.73 ± 0.09, while the GCN-based model achieved an NRMSE of 9.47 ± 1.77% and an R^2 of 0.71 ± 0.13; the two models thus performed comparably. The proposed EdgeConv-based model, by contrast, achieved an NRMSE of 8.79 ± 1.26% and an R^2 of 0.76 ± 0.09, showing a lower NRMSE and a higher R^2 than the other two models.

Figure 3 shows the difference in the NRMSE for each model relative to the Transformer-based model (ΔNRMSE). For F_x, the NRMSE decreased in six of the 12 participants with the GCN-based model and in seven of the 12 participants with the EdgeConv-based model. The EdgeConv-based model showed an overall improvement relative to the Transformer-based model, with a mean ΔNRMSE of −0.51 pt (median −0.32 pt), whereas the GCN-based model showed little change (mean +0.16 pt, median +0.17 pt). For F_z, by contrast, the GCN-based model showed a greater improvement than the EdgeConv-based model, with ΔNRMSE values of −0.41 pt and +0.13 pt, respectively.

To examine the effect of EdgeConv across the gait cycle, each stance phase stride was divided into four sub-phases: loading response (LR), mid-stance (MSt), terminal stance (TSt), and pre-swing (PSw). The mean absolute error (MAE) for F_x was compared across models within each sub-phase (Figure 4). The EdgeConv-based model showed an error similar to that of the Transformer-based model in LR, a slightly lower error in MSt, and a lower MAE than the Transformer-based model in both TSt and PSw, with the difference between the two models being comparatively large in TSt and PSw. The GCN-based model did not exhibit a clear phase-dependent error reduction, as observed for the EdgeConv-based model.

To examine the contribution of the spatial relationships used by the EdgeConv-based model, an ablation analysis was performed on the trained model, in which each edge was removed individually (Figure 5). For F_x, removing the edge between the 3rd and 4th metatarsal heads produced the greatest increase in the NRMSE (ΔNRMSE ≈ +1.58 pt); removing the edges between the hallux and the 1st metatarsal head, between the lateral and medial heel, and between the lateral and central heel also produced increases (+0.70 pt, +0.61 pt, and +0.28 pt, respectively). By contrast, removing the edge between the 1st and 3rd metatarsal heads or between the medial and central heels produced little change, and removing the edge between the lateral arch and lateral heel produced a slight reduction in the NRMSE. These results indicate that not all edges used by EdgeConv contribute equally and that specific inter-sensor relationships have a substantially greater influence on F_x estimation than others.

Regarding comparison with prior work, a previous study that estimated GRF using a many-to-one LSTM model trained separately for each axis from stance-phase plantar pressure insole data reported an NRMSE of 12.96 ± 4.98% for the mediolateral component. The NRMSE for F_x achieved by the proposed EdgeConv-based model in this study (8.79 ± 1.26%) was lower than this previously reported value. Nonetheless, this comparison should be interpreted with caution, as the prior study employed a many-to-one regression model trained independently for each axis, whereas we employed a many-to-many regression model that simultaneously estimated F_x, F_y, and F_z. Therefore, the two studies differed in terms of model architecture and evaluation conditions.

Table 3: Estimation accuracy (NRMSE) of the three models in the between-subject evaluation.

Model	*Fx*	F_y	F_z
Transformer-based model	9.30 ± 1.52	5.84 ± 1.83	5.37 ± 1.34
GCN-based model	9.47 ± 1.77	5.75 ± 2.07	4.96 ± 1.33
EdgeConv-based model	8.79 ± 1.26	5.88 ± 2.55	5.50 ± 1.77

Table 4: Estimation accuracy (R^2) of the three models in the between-subject evaluation.

Model	*Fx*	*Fy*	*Fz*
Transformer-based model	0.73 ± 0.09	0.90 ± 0.07	0.94 ± 0.03
GCN-based model	0.71 ± 0.13	0.90 ± 0.09	0.95 ± 0.03
EdgeConv-based model	0.76 ± 0.09	0.89 ± 0.12	0.94 ± 0.04

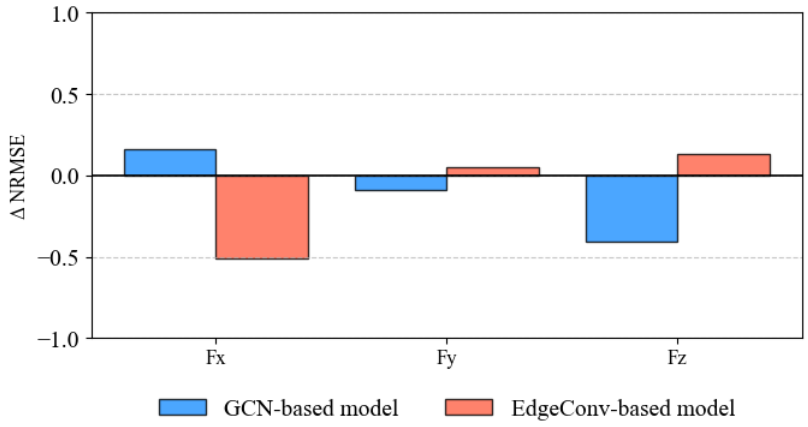


Figure 3: Changes in the normalized root mean square error (NRMSE) relative to the Transformer-based model.

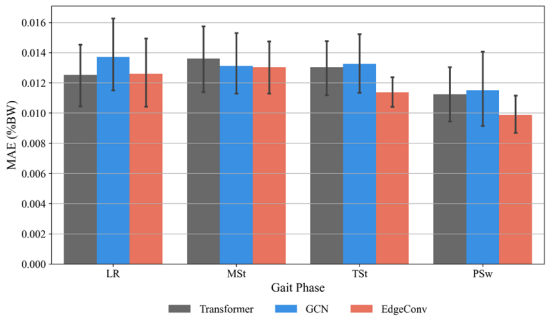


Figure 4: Comparison of mean absolute error (MAE) for F_x across gait sub-phases.

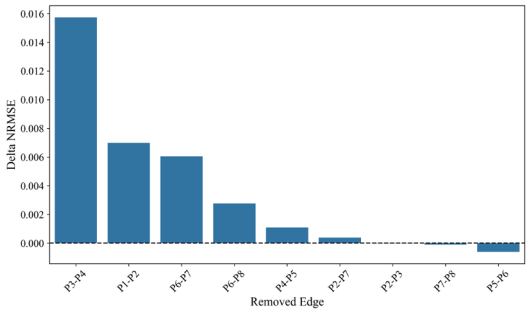


Figure 5: Effect of individual edge removal on normalized root mean square error (NRMSE) for F_x estimation.

DISCUSSION

The individual-participant comparison for F_x showed that, for the GCN-based model, a largely equal number of participants showed improved and worsened NRMSE relative to the Transformer-based model, whereas for the EdgeConv-based model, more participants showed improvement, and both the mean and median ΔNRMSE were in the direction of improvement. This result suggests that simply introducing a graph structure does not necessarily improve the F_x estimation accuracy and that the mechanism whereby EdgeConv explicitly learns local difference features may have produced a modest improvement trend. In a GCN, the process of aggregating neighboring node features may smooth out local spatial information, whereas in EdgeConv, the difference between the central node and its neighboring nodes is explicitly treated as a feature, making it less likely for local pressure-change information to be lost.

The phase-resolved error analysis showed that the improvement from EdgeConv was not uniform across the stance phase but was instead concentrated in the terminal stance and pre-swing phases. This portion of the stance involves the foot rolling forward and the load shifting toward the forefoot, with active load transfer around the metatarsal heads and hallux and correspondingly greater changes in the mediolateral pressure distribution within the sole. The local pressure-difference features learned by EdgeConv may have captured this mediolateral load-transfer information, contributing particularly to the F_x estimation during this period.

The edge ablation results were consistent with this interpretation. The edges that contributed the most to the F_x estimation were all connected sensors that were adjacent in the mediolateral (width) direction of the foot, with the connections between the metatarsal heads in the forefoot and between the lateral and medial heels being particularly important. These edges are well positioned to directly capture the mediolateral load transfer that F_x reflects, and the importance of the forefoot edges is consistent with the improvement observed in the terminal stance and pre-swing in the phase-resolved analysis. In contrast, the edge added in this study to span the medial longitudinal arch showed almost no effect when removed, indicating that its contribution to F_x estimation was limited, contrary to our expectations.

No clear improvement from EdgeConv was observed for F_y or F_z; in particular, for F_z, the GCN-based model achieved a lower NRMSE than the EdgeConv-based model. As F_z is strongly associated with the total plantar load, the additional benefit of introducing local difference features may have been limited to this component, although this remains speculative.

CONCLUSION

This study developed an EdgeConv-based model that learns the spatial difference features of plantar pressure. In the evaluation using subject-level CV, the proposed model achieved an NRMSE of 8.79 ± 1.26% and an R^2 of 0.76 ± 0.09 for F_x, outperforming both the Transformer- and GCN-based models. The individual-participant comparison also showed a more consistent improvement trend for EdgeConv than for GCN. These findings suggest that explicitly learning local spatial differences in plantar pressure may provide a useful basis for estimating mediolateral GRFs from wearable in-shoe sensor data, potentially contributing to continuous gait assessment outside laboratory environments.

REFERENCES

Bauby, C.E., Kuo, A.D. 2000. "Active control of lateral balance in human walking", Journal of Biomechanics, 33 (11), 1433–1440.

Bowden, M.G., Balasubramanian, C.K., Neptune, R.R., Kautz, S.A. 2006. "Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking", Stroke, 37 (3), 872–876.

Hajizadeh, M., Clouthier, A.L., Kendall, M., Graham, R.B. 2023. "Predicting vertical and shear ground reaction forces during walking and jogging using wearable plantar pressure insoles", Gait & Posture, 104, 90–96.

Kim, J., Kang, H., Lee, S., Choi, J., Tack, G. 2023. "A deep learning model for 3D ground reaction force estimation using shoes with three uniaxial load cells", Sensors, 23 (7), 3428.

Ministry of Health, Labor and Welfare, Government of Japan. 2025. Comprehensive Survey of Living Conditions. Website: https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa25/dl/05.pdf

Shah, V.V., McNames, J., Mancini, M., et al., 2020. "Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls", Journal of NeuroEngineering and Rehabilitation, 17, Article 159.

Xia, Z., Sun, D., Qian, Y., et al., 2026. "Biomechanical strategies for minimizing force plate targeting effects during running: Efficacy of masked force plate integration with augmented visual feedback", Journal of Human Kinetics, 102, 5–12.