

Original Paper

Autoencoder-Enhanced Convolutional Neural Networks for Plantar Pressure–Based Gait Pattern Recognition: Model Development and Cross-Validated Evaluation Study

Chuan-Chun Chang¹, MS; Chi-Wen Lung^{2,3}, PhD; Yih-Kuen Jan³, PhD; Qi-Qian Lu¹, MS; Yi-You Wang¹, MS; Yi-Sheng Chen⁴, PhD; Ben-Yi Liao¹, PhD

¹Department of Automatic Control Engineering, Feng Chia University, Taichung, Taiwan

²Department of Creative Product Design, Asia University, Taichung, Taiwan

³Rehabilitation Engineering Lab, Department of Health and Kinesiology, University of Illinois Urbana-Champaign, Urbana, IL, United States

⁴Department of Communications Engineering, Feng Chia University, Taichung, Taiwan

Corresponding Author:

Ben-Yi Liao, PhD

Department of Automatic Control Engineering

Feng Chia University

No. 100 Wenhua Road, Xitun District

Taichung, 407102

Taiwan

Phone: 886 4 24517250 ext 3915

Email: byliao@fcu.edu.tw

Abstract

Background: Plantar pressure imaging is a stable modality that reflects gait-related biomechanical characteristics and has been used increasingly for gait assessment and recognition. However, plantar pressure images are high dimensional and nonlinear, making manual feature engineering and conventional machine learning insufficient to capture discriminative patterns.

Objective: This study aimed to develop a gait pattern recognition model based on plantar pressure using an autoencoder (AE)-enhanced convolutional neural network (CNN) and to evaluate its performance against baseline deep learning and classical machine learning approaches.

Methods: A total of 13 healthy volunteers (aged 18-24 years) were recruited. Plantar pressure data were collected during treadmill walking using an in-shoe pressure measurement system and converted into frame-wise plantar pressure images. We compared a lightweight CNN (Light CNN), an AE-CNN cascade model, and an encoder-augmented CNN with an additional bottleneck layer. Model development used participant-wise data partitioning, and performance was evaluated using accuracy, precision, recall, and F_1 -score.

Results: The proposed encoder-augmented CNN achieved the best overall performance (F_1 -score=96.20%), outperforming the Light CNN (F_1 -score=94.44%) and AE-CNN cascade (F_1 -score=92.45%). Confusion matrices and learning curves further indicated stable training behavior and consistent classification performance across gait patterns.

Conclusions: Integrating representation learning (AE-based compression) with CNN-based classification improved the recognition of gait patterns from plantar pressure images. This pilot study included only healthy participants. Future work should validate generalizability in larger and clinically diverse cohorts and further investigate participant-level evaluation and model interpretability, as well as deployment feasibility.

(JMIR Form Res 2026;10:e88488) doi: [10.2196/88488](https://doi.org/10.2196/88488)

KEYWORDS

plantar pressure; gait; convolutional neural networks; autoencoder; wearable sensors; deep learning; machine learning; biomedical engineering

Introduction

With the rapid development of smart health care and artificial intelligence, the analysis of physiological signals has become increasingly important in fields such as disease diagnosis, rehabilitation monitoring, and identity verification. Among these signals, plantar pressure imaging is a stable method that reflects the posture of the individual, the characteristics of the gait, and foot functionality. Due to its ease of collection and visualization, plantar pressure imaging has been widely used for behavioral analysis, clinical evaluation, and biometric recognition. Gait is a fundamental human activity that represents the coordinated movement of the lower limbs. Although often overlooked because of its automaticity, gait involves complex and synchronized interactions between the musculoskeletal system and neural control. Understanding gait dynamics is important to assess gait alterations and support rehabilitation planning. Previous studies have investigated gait analysis using deep spatiotemporal learning and machine learning, including pathological gait classification, gait phase detection, pilot gait classification studies, and deep learning-based gait trajectory modeling [1-4]. Typically, the gait cycle can be described by the stance and swing phases, and more detailed phase characterization may provide additional functional information [2,4].

Traditional gait and plantar pressure analysis relied on manual feature extraction and classical machine learning algorithms. However, plantar pressure data are typically high dimensional and nonlinear, making it difficult for handcrafted features to capture informative patterns consistently. In recent years, deep learning, particularly convolutional neural networks (CNNs), has shown superior performance in automatically extracting features and improving classification accuracy. Autoencoder (AE) techniques have also been used to reduce dimensionality and alleviate redundancy, improving representation compactness and computational efficiency for high-dimensional inputs. Previous research has explored plantar pressure-based assessment in health-related applications. Deschamps et al [5] analyzed plantar pressure distribution patterns in people with diabetes, and Amemiya et al [6] investigated the relationships between elevated plantar pressure and gait characteristics in patients with diabetes. Wang et al [7] developed an insole-based gait monitoring technique to recognize gait patterns associated with knee osteoarthritis. These studies highlight the value of plantar pressure analysis to quantify lower limb loading characteristics and support gait-related assessment. Other researchers have investigated computational models for gait recognition. Nguyen et al [8] used smart shoes to classify ambulatory activities and proposed statistical characteristics combined with conventional classifiers, and Jeong et al [9] studied the classification of activity with respect to stairs using plantar pressure sensors. Jun et al [10] further proposed a hybrid

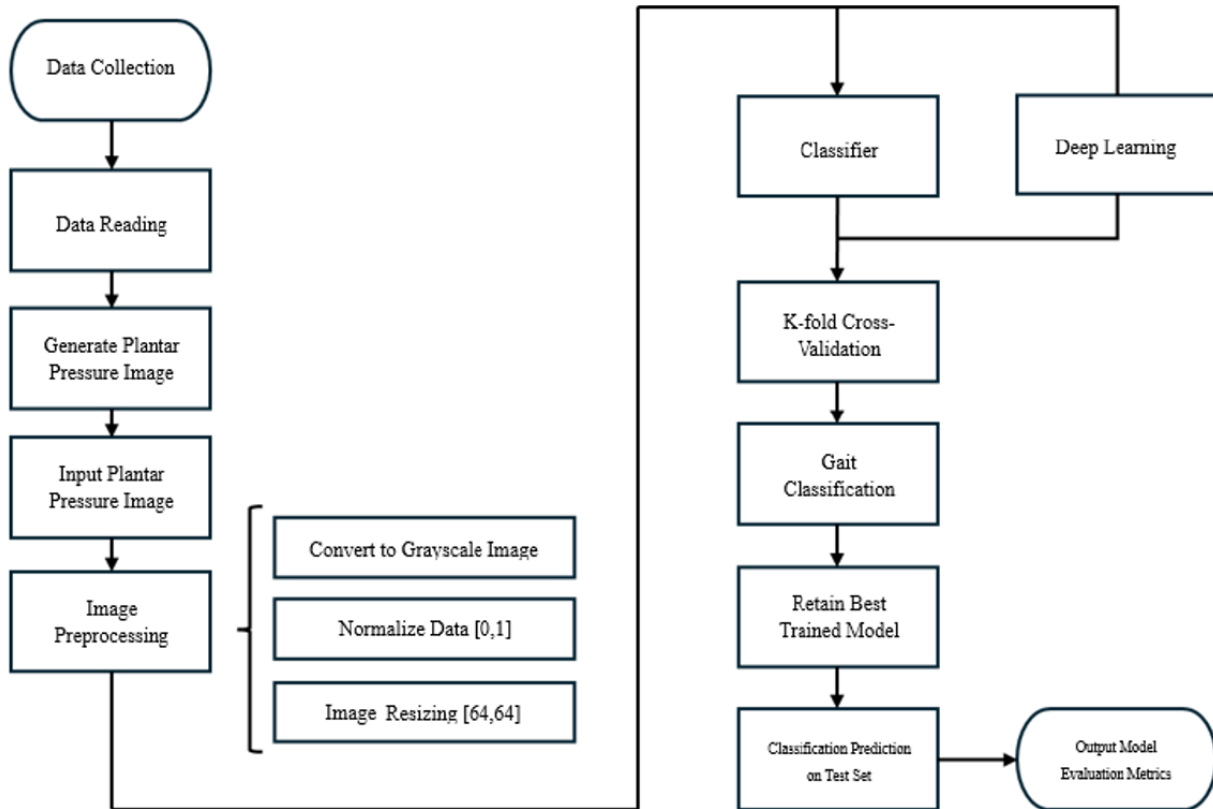
deep learning framework that integrates plantar pressure images and 3D skeletal data, indicating that multimodal fusion can improve the recognition of abnormal gait patterns compared to pressure-only inputs. Beyond classification, related gait evaluation studies have compared optimization strategies for sensory data classification using deep neural networks [11] and evaluated machine learning algorithms for electromyography pattern classification in gait disorders [12], underscoring the broader demand for robust learning pipelines. To address the need for portable, real-time applications, Cho [13] developed a deep learning approach using plantar pressure signals to estimate walking speed and gait-related classification tasks. Chhoeum et al [14] applied CNN-based regression to estimate knee joint angles using foot pressure mapping images. Ling et al [15] introduced an AE-CNN-based multisource data fusion framework to estimate the step length of the gait motion, illustrating the practical value of representation learning when handling high-dimensional gait-related data. Ardhiyanto et al [16] formulated the estimation of the foot progression angle as an object detection task on plantar pressure images using YOLO (You Only Look Once)-based models, demonstrating the feasibility of computer vision-style pipelines on pressure maps. From a system-level perspective, Zhou et al [17] developed a gait detection and plantar pressure analysis system using a flexible triboelectric pressure sensor array and deep learning, highlighting the direction of continuous wearable gait monitoring under real-world constraints. Collectively, these studies suggest that plantar pressure imaging is advantageous due to its stability, visualization quality, and strong correspondence with gait cycles. Building on these advances, there remains a need for pressure-only deep learning frameworks that can learn discriminative representations from high-dimensional plantar pressure images, reduce redundancy via representation learning (eg, AEs), and maintain computational efficiency. Importantly, clinical claims should be supported by validation in clinically diverse cohorts; therefore, model development and evaluation in controlled settings should be clearly distinguished from future clinical verification.

Methods

Study Design and Workflow

The general workflow of this study is illustrated in Figure 1. The experimental design consisted of 3 major components. The first component involved the collection of plantar pressure response data required to generate plantar pressure images. The second component focused on the application of machine learning classifiers and a deep learning model based on a CNN. The third component addressed the AE, which was further investigated as a core part of this study. In the final stage of the experiments, the CNN and AE frameworks were integrated to perform classification and performance evaluation.

Figure 1. Overall workflow of the proposed plantar pressure–based recognition system.



As shown in Figure 1, the experimental workflow highlights the sequential architecture of the study. In particular, the second component, which encompasses both classifiers and CNN-based

models, is presented in separate blocks. The detailed structures of these blocks are illustrated in Figures 2 and 3.

Figure 2. Model development framework that includes the autoencoder (AE)–convolutional neural network (CNN) pipeline and traditional classifiers. AE: autoencoder; AE-CNN: autoencoder convolutional neural network; KNN: k-nearest neighbors; SVM: support vector machine.

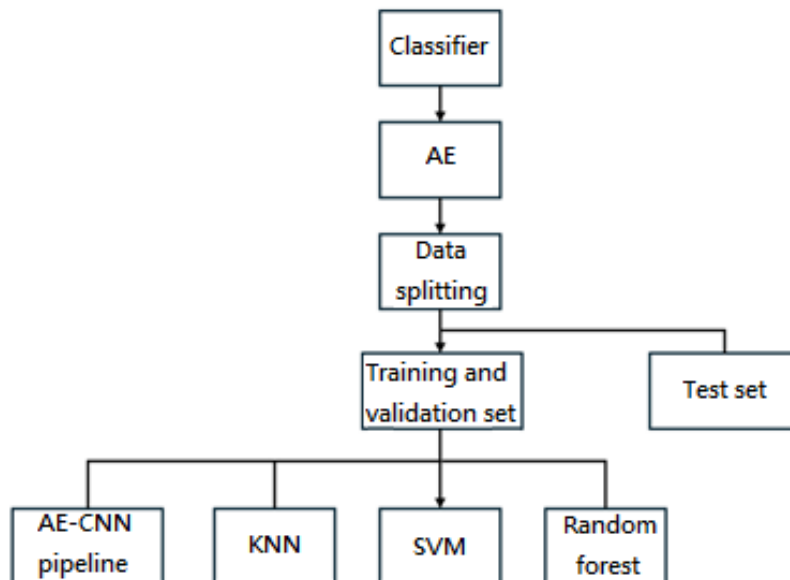
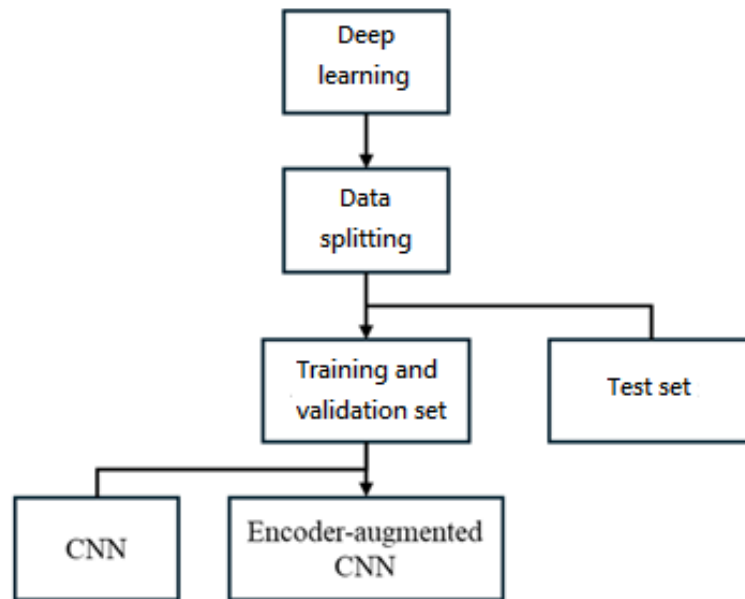


Figure 3. Architectures of the convolutional neural network (CNN) and encoder-augmented CNN models.



In the classifier block diagram, the AE-CNN cascade is defined as an AE in which compressed data are modularized and sequentially integrated with a CNN. This design enables the model to be trained in a staged manner, thereby forming the AE-CNN cascade. Furthermore, because its data processing pathway resembles that of conventional classifiers, the AE-CNN cascade was grouped within the classifier block for consistency in experimental design.

In the deep learning block diagram, the encoder-augmented CNN is defined as a hybrid model in which the encoder component of the AE is directly integrated into the CNN classifier. This design enables end-to-end training, allowing the encoder and CNN to fuse into a unified framework. By adopting this hybrid strategy, the model is trained as an integrated architecture, which is referred to as encoder-augmented CNN.

Participants and Experimental Setup

A total of 13 healthy university student volunteers (aged 18-24 years) were recruited for this study. Exclusion criteria included current or previous foot ulcers, diabetes, vascular disease, hypertension, inability to walk independently for at least 10 minutes, and continued use of medications that could affect

gait. Only participants who met all eligibility criteria were enrolled.

Ethical Considerations

This study was approved by the Central Regional Research Ethics Committee of China Medical University, Taichung, Taiwan (approval CRREC-112-130). All participants received a full explanation of the study procedures and provided their written informed consent prior to participation. The data were deidentified for analysis and reporting. No financial compensation was provided to participants in this study.

Plantar Pressure Measurement

Plantar pressure data were acquired using the Tekscan F-Scan in-shoe pressure measurement system (Figure 4 [18]). The F-Scan provides high-resolution plantar pressure distribution with real-time acquisition via a dedicated data cable, making it suitable for gait analysis, sports science, insole design, and clinical applications. The key functions adopted in this study are summarized in Table 1. In particular, the system's real-time recording capability and high spatial resolution enabled the capture of subtle pressure changes, thereby improving the fidelity of the experimental dataset.

Figure 4. F-Scan plantar pressure sensor: (A) data collection wearing mode and (B) pressure sensing sheet [18].



Table 1. Main functions of the F-Scan plantar pressure sensor.

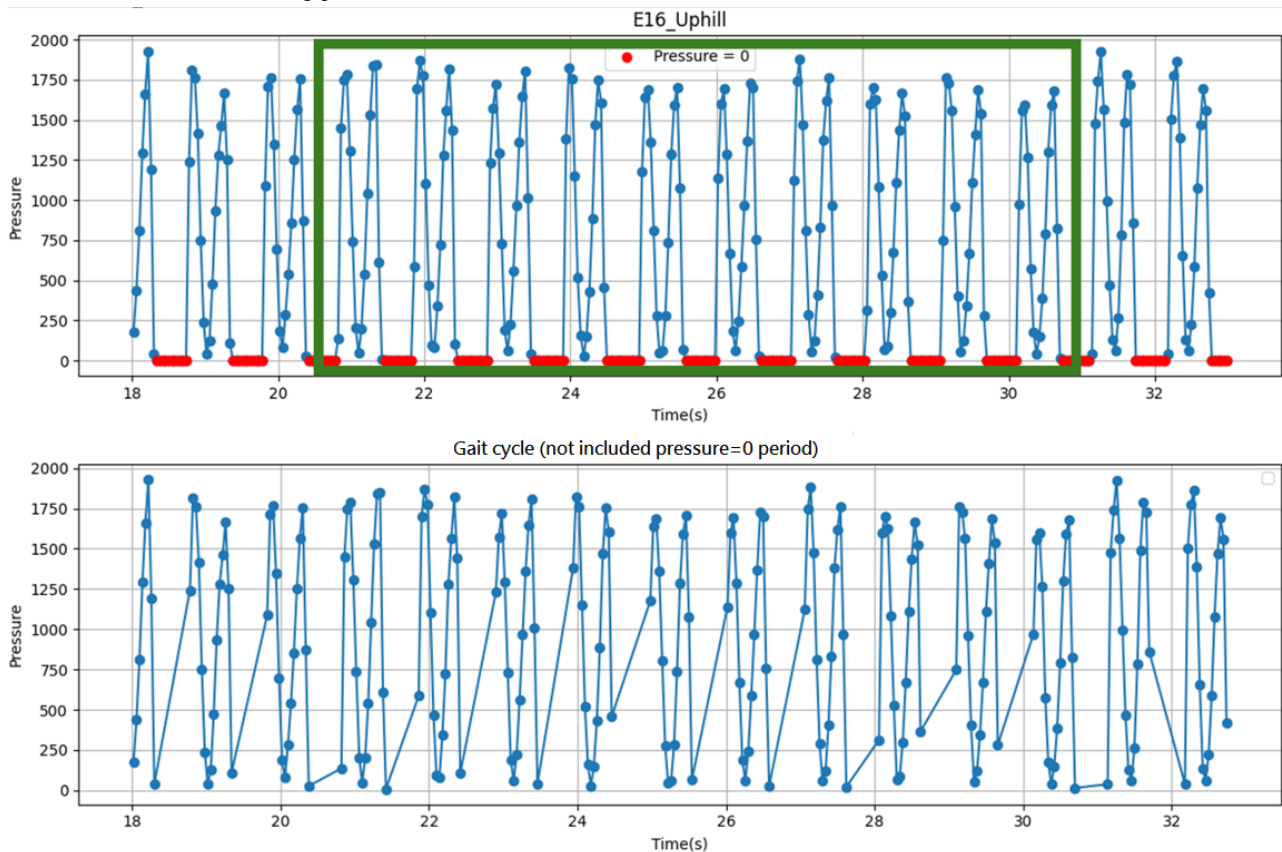
Item	Explanation	Application
Plantar pressure measurement	Unit: kPa	Measurement of plantar pressure distribution
Gait analysis	Analysis of step length, step frequency, and gait cycle	Walking, running, and abnormal gait
High-resolution sensor	0.5 cm ² per sensing unit; visualization of subtle pressure variations	Detailed plantar pressure analysis
Calibration function	Three calibration methods: walking, gait, and point modes	Normal gait analysis, abnormal gait analysis, and plantar pressure distribution while standing

Foot Pressure Data Acquisition

Plantar pressure responses were recorded using the Tekscan F-Scan in-shoe system while participants walked on a treadmill at a fixed speed. Each trial began with a 1-minute familiarization period to stabilize gait at the target speed. The F-Scan microsensors sampled plantar pressure at 25 Hz, and instantaneous pressures were stored as matrix-valued frames and exported in CSV format for downstream processing.

Frame selection was based on gait cycles, each consisting of a stance phase and a swing phase (Figure 5). The pressure value at each time point was calculated as the sum of pressure readings across all sensing elements (kPa) as an overall loading indicator. In Figure 5, the green bracket denotes 10 gait cycles, blue markers indicate frames within the stance phase, and red markers denote the swing phase. As the plantar load is negligible during swing, the sensors yield near-zero readings. Therefore, to enable frame-wise classification of gait patterns and only the stance-phase frames of the 10 cycles were retained for analysis.

Figure 5. Example of the plantar pressure response across gait cycles: the green bracket indicates 10 gait cycles, blue markers denote stance-phase frames, and red markers denote swing-phase frames.



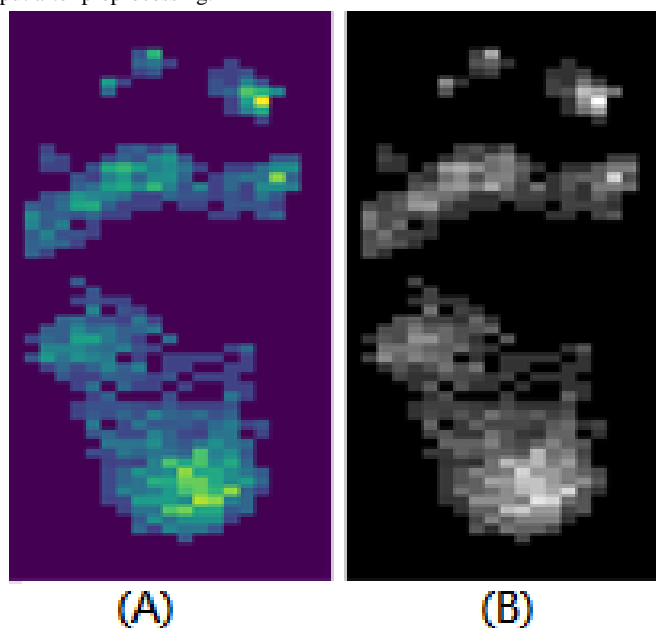
The final dataset comprised 6994 frames in 3 gait patterns: slow walking (2590 frames), fast walking (2162 frames), and uphill walking (2242 frames). These frames served as input to the proposed models for training and evaluation.

Image Preprocessing

Matrix-form plantar pressure signals exported by the F-Scan system were converted into frame-wise instantaneous plantar pressure matrices using Python 3.10.12 (Python Software Foundation) to enable deep learning. Each frame-wise matrix was normalized to the range (0, 1) using min-max normalization

and rendered directly as a grayscale intensity image, in which pixel intensity represents normalized pressure magnitude. For visualization (Figure 6A), we additionally rendered the same normalized matrices as pseudocolor pressure maps using the perceptually uniform viridis colormap. The viridis colormap was used for visualization only. The grayscale images were resized to 64×64 pixels to ensure consistency across samples, and Gaussian noise was injected as a lightweight augmentation to improve noise robustness and mitigate overfitting. After preprocessing, the resulting input tensor for model training and evaluation had the shape (6994, 64, 64, 1).

Figure 6. (A) Example of an image of plantar pressure distribution rendered with a perceptually uniform viridis colormap. (B) The corresponding grayscale image used as the model input after preprocessing.



Model Architecture

In this study, 3 deep learning architectures were developed to address plantar pressure image classification, each designed with distinct structural characteristics. The first was a lightweight CNN (Light CNN), which served as the baseline model. It was constructed with 3 convolutional layers, each followed by max-pooling to progressively reduce dimensionality

while preserving salient spatial features. Batch normalization was incorporated to stabilize the training process, while dropout layers were used to mitigate overfitting. A fully connected layer and a final softmax classifier were used to output the probabilities of the 3 gait categories, namely, slow walking, fast walking, and uphill walking. The structural design of this network is summarized in Table 2, which establishes the baseline framework for subsequent comparisons.

Table 2. Light convolutional neural network model architecture.

Tasks	Structural parameters		Training parameters (activation function)
	Layer	Detailed parameters	
Feature extraction	Input image	(64, 64, 1)	— ^a
Feature extraction	Conv2D_1	32,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	Leaky ReLU
Feature extraction	MaxPool2D	(2, 2), padding="same"	—
Feature extraction	Conv2D_2	64,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	Leaky ReLU
Feature extraction	MaxPool2D	(2, 2), padding="same"	—
Feature extraction	Conv2D_3	128,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	Leaky ReLU
Feature extraction	MaxPool2D	(2, 2), padding="same"	—
Feature extraction	Dropout (0.2)	—	—
Multiclass classification	Dense	90	—
Multiclass classification	Dense	60	—
Multiclass classification	Dense	3	Softmax
Other parameters	K-fold	9	—
Other parameters	Batch size	64	—
Other parameters	Epochs	100	—
Other parameters	Loss function	categorical_crossentropy	—
Other parameters	Optimizer	Adam (learning rate=0.0005)	—
Other parameters	Metrics	accuracy	—

^aNot applicable.

The second model was the AE-CNN cascade, which modularized the integration of an AE and a CNN classifier. In this design, the encoder of the AE first compressed the high-dimensional plantar pressure matrices into a compact latent representation, while the decoder simultaneously ensured that essential structural information could be reconstructed. The compressed latent features were then transferred to a CNN

classifier for further convolutional processing and final classification. This staged cascade design effectively separated feature compression from classification, improving both the stability and the interpretability of the learning process. The encoder architecture used in the AE-CNN cascade is presented in [Table 3](#), illustrating how feature reduction and classification were sequentially integrated.

Table 3. Autoencoder encoder model architecture.

Task	Structural parameters		Training parameters (activation function)
	Layer	Detailed parameters	
Encoded	Input image	(64, 64, 1)	— ^a
Encoded	Conv2D_1	32, (3, 3), padding="same"	Leaky ReLU
Encoded	MaxPool2D	(2, 2), padding="same"	—
Encoded	Conv2D_2	64, (3, 3), padding="same"	Leaky ReLU
Encoded	MaxPool2D	(2, 2), padding="same"	—
Encoded	Conv2D_3	128, (3,3), padding="same"	Leaky ReLU
Encoded	MaxPool2D	(2, 2), padding="same"	—
Encoded	Flatten	—	—
Encoded	Dense	128	—

^aNot applicable.

To evaluate generalization while reducing the risk of information leakage from correlated frame-wise samples, data partitioning was performed at the participant level. Participants ($n=13$) were first split into a development set ($n=10$, 76.9%) and an independent held-out test set ($n=3$, 23.1%). Within the development set, a participant-wise grouped 9-fold cross-validation procedure was applied for model selection and stability assessment. In each fold, all frame-wise samples from participants in the training folds were used for model fitting, and all samples from participants in the held-out fold were used for validation. As 10 is not evenly divisible by 9, fold sizes differed by at most 1 participant (8 folds contained 1 participant and 1 fold contained 2 participants). After cross-validation, the final model configuration was retrained on the full development

set and evaluated once on the held-out test set. Performance was reported using accuracy, precision, recall, and F_1 -score.

The third architecture was the encoder-augmented CNN, which further extended the integration of AE and CNN by embedding the encoder directly into the CNN pipeline. Unlike the cascade structure, this hybrid design adopted an end-to-end framework, where the encoder served as the initial feature extractor and its outputs were directly connected to the subsequent CNN layers. This approach allowed the encoder and CNN to be jointly optimized, combining compact representation learning with the discriminative power of deep convolutional layers. The architectural layout of this model is summarized in Table 4, highlighting its streamlined structure and enhanced learning efficiency.

Table 4. Encoder-augmented convolutional neural network model architecture.

Task	Structural parameters		Training parameters (activation function)
	Layer	Detailed parameters	
Spatial and feature compression	Input image	(64, 64, 1)	— ^a
Spatial and feature compression	Conv2D_1	32,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	—
Spatial and feature compression	Batch normalization	—	Leaky ReLU
Spatial and feature compression	MaxPool2D	(2, 2), padding="same"	—
Spatial and feature compression	Conv2D_2	64,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	Leaky ReLU
Spatial and feature compression	MaxPool2D	(2, 2), padding="same"	—
Spatial and feature compression	Conv2D_3	128,(3, 3), padding="same"; L2 regularization ($\lambda=0.0005$)	Leaky ReLU
Spatial and feature compression	MaxPool2D	(2, 2), padding="same"	—
Spatial and feature compression	Dropout(0.2)	—	—
Spatial and feature compression	Flatten	—	—
Spatial and feature compression	Dense (bottleneck)	128	—
Multiclass classification	Dense	90	—
Multiclass classification	Dense	60	—
Multiclass classification	Dense	3	Softmax
Other parameters	K-fold,	9	—
Other parameters	Batch size	64	—
Other parameters	Epochs	100	—
Other parameters	Loss function	categorical_crossentropy	—
Other parameters	Optimizer	Adam (learning rate=0.0005)	—
Other parameters	Metrics	accuracy	—

^aNot applicable.

Evaluation Metrics

Model performance was assessed using accuracy, precision, recall (sensitivity), F_1 -score, and confusion matrix:

- The accuracy reflects the overall proportion of correctly classified samples in all classes.

- Precision quantifies the reliability of positive predictions for a given class by indicating how many predicted positives are correct.
- Recall (sensitivity) measures the model's ability to correctly identify true instances of a given class.

- The F_1 -score provides a balanced summary of precision and recall, particularly useful when both false positives and false negatives matter.
- The confusion matrix offers a class-by-class view of predictions vs ground truth, enabling identification of error patterns (eg, which gait classes are most frequently confused). For the 3-class setting in this study (slow walking, fast walking, and uphill walking), a 3×3 matrix was used to summarize the results per class and general trends.

All metrics were computed on the designated evaluation split without using any evaluation data for parameter updating. Unless stated otherwise, results are reported at the overall level to facilitate comparison among the baseline CNN, the AE-CNN cascade, the encoder-augmented CNN, and traditional classifiers; confusion matrices are additionally provided to illustrate class-wise error patterns. To reduce optimistic bias due to correlated frame-wise samples, all data splitting was performed at the participant level. Specifically, all stance-phase frames from the same participant were assigned to a single fold during cross-validation, ensuring that no participant's frames appeared in both the training and validation sets within any iteration. Model performance was summarized using accuracy, precision, recall, and F_1 -score.

Reporting Guidelines

This model development and evaluation study was reported with reference to the CREMLS (Consolidated Reporting of Machine Learning Studies) checklist. The completed checklist is provided in [Multimedia Appendix 1](#).

Results

Overview

This section summarizes the main comparative findings from the deep learning and classical machine learning models evaluated. Detailed training curves, optimization experiments, and hyperparameter tuning results are provided in [Multimedia Appendix 2](#).

Among the baseline deep learning models, the Light CNN achieved an F_1 -score of 94.44% on the held-out test set, whereas the AE-CNN cascade achieved an F_1 -score of 92.45%.

Additional optimization analyses of the encoder-augmented CNN, including comparisons of downsampling strategies, batch normalization configurations, and bottleneck layer inclusion, supported the final model design reported here ([Multimedia Appendix 2](#)).

For the classical machine learning models trained on AE-derived features, support vector machine (SVM) with a radial basis function kernel performed best (F_1 -score=93.76%), followed by k-nearest neighbors (F_1 -score=91.73%) and random forest (F_1 -score=88.54%).

Principal Findings

The primary finding of this study is that the proposed encoder-augmented CNN architecture achieves superior performance (F_1 -score=96.20%) in classifying dynamic gait patterns from plantar pressure images compared to both baseline deep learning models and classical machine learning classifiers ([Figure 7](#); [Table 5](#)). This result supported the hypothesis that an integrated deep learning architecture, combining the feature extraction and dimensionality reduction capabilities of an AE with the classification power of a CNN, is highly effective for this task.

Figure 7. Overall performance comparison of all models. AE-CNN: autoencoder convolutional neural network; KNN: k-nearest neighbors; SVM-RBF: support vector machine-radial basis function.

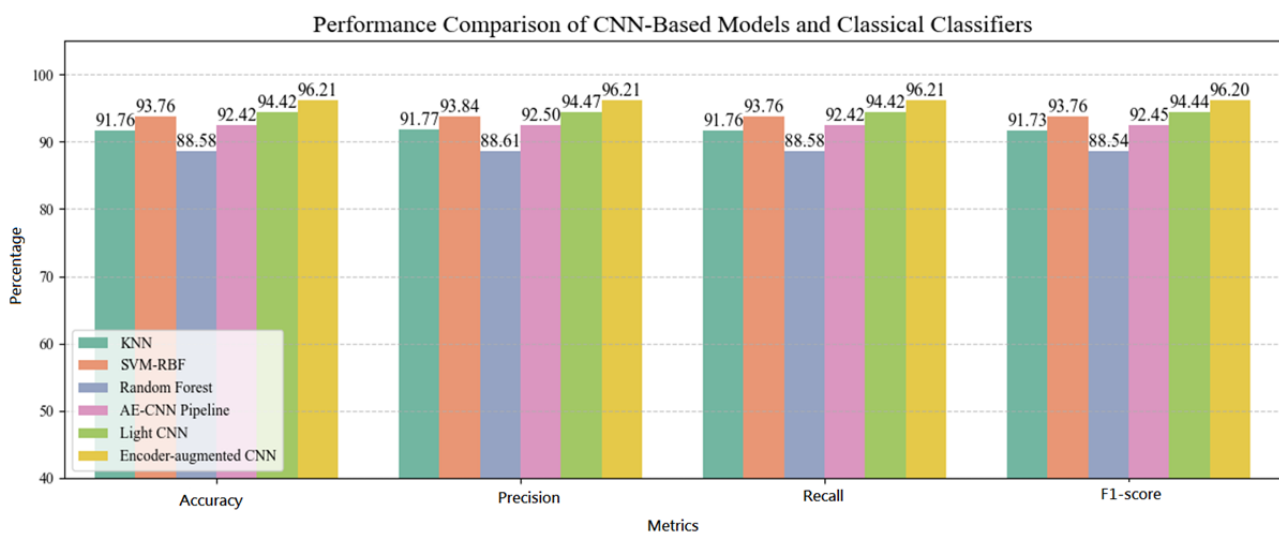


Table 5. Overall performance in the held-out test set.

Model	Accuracy	Precision	Recall	F_1 -score
KNN ^a	91.76	91.77	91.76	91.73
SVM ^b -RBF ^c	93.76	93.84	93.76	93.76
Random forest	88.58	88.61	88.58	88.54
AE ^d -CNN ^e cascade	92.42	92.50	92.42	92.45
Light CNN	94.42	94.47	94.42	94.44
Encoder-augmented CNN	96.21	96.21	96.21	96.20

^aKNN: k-nearest neighbors.

^bSVM: support vector machine.

^cRBF: radial basis function.

^dAE: autoencoder.

^eCNN: convolutional neural network.

Discussion

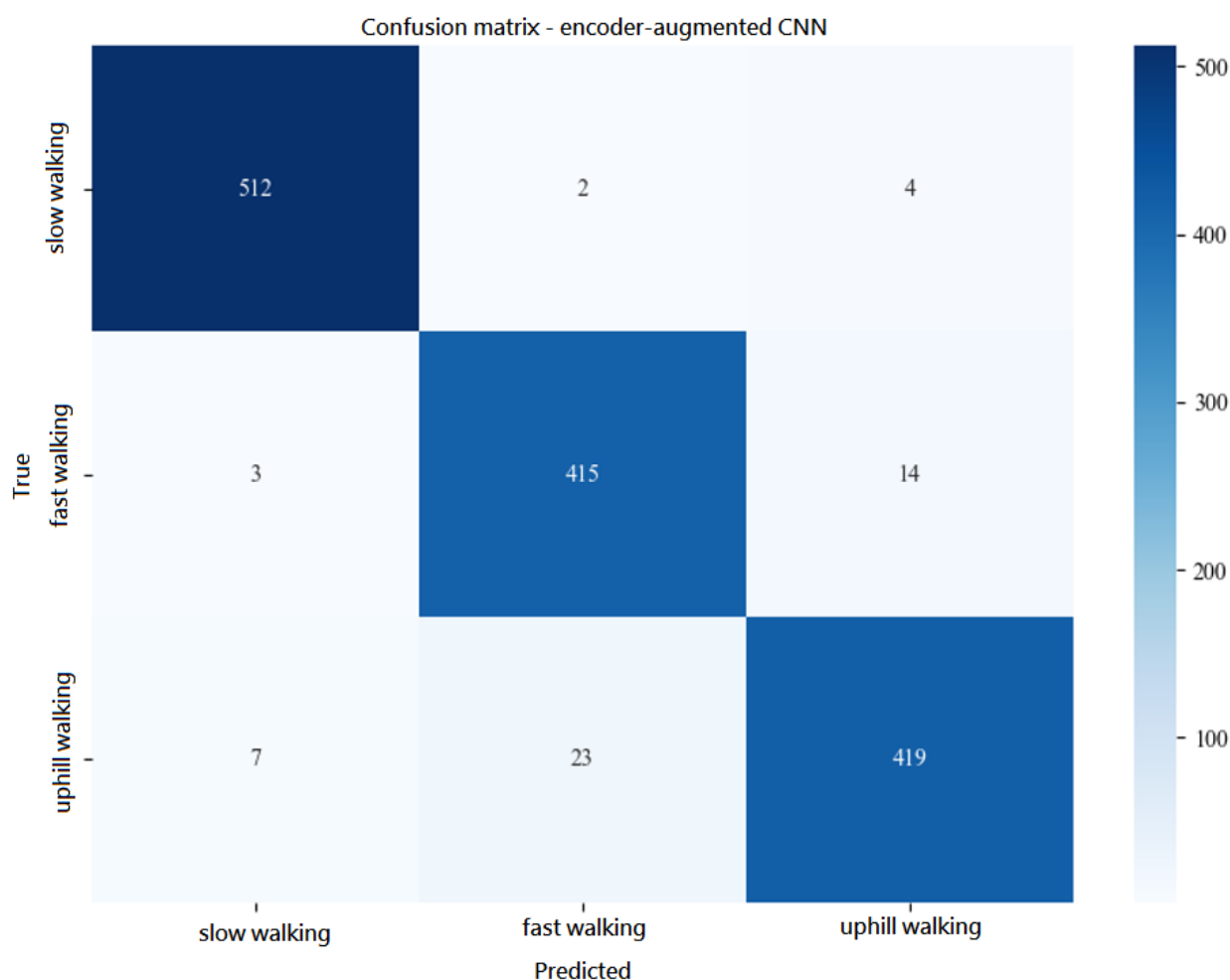
Overview

This model development and evaluation study demonstrates that the integration of AE-based representation learning with CNN architectures can achieve accurate recognition of gait patterns based on plantar pressure in a pilot dataset of healthy participants. Future work should validate generalization across broader demographics and clinical populations, evaluate robustness to real-world variability (eg, footwear, sensor noise, and speed fluctuations), and further strengthen interpretability and deployment feasibility for wearable or embedded applications. The performance of the Light CNN (F_1 -score=94.44%) also demonstrated that even a lightweight, stand-alone CNN can effectively learn discriminative features from plantar pressure data. Among the classical methods, the **support vector machine**–RBF classifier proved to be the most robust, outperforming **k-nearest neighbors** and random forest, which aligns with its known strengths in handling high-dimensional feature spaces.

The study also highlights the importance of architectural choices in model optimization. Supplementary analyses of downsampling strategy, batch normalization placement, and bottleneck layer inclusion are provided in [Multimedia Appendix 2](#). These experiments supported the final selection of the encoder-augmented CNN configuration reported in the main manuscript.

Interpretation of Results and Gait Feature Analysis

A notable and consistent finding in multiple models was the confusion between “fast walking” and “uphill walking” gaits ([Figure 8](#)). The encoder-augmented CNN, for example, misclassified 23 “uphill” instances as “fast walking.” This suggests that at similar walking speeds, the plantar pressure distributions for these 2 activities share substantial similarities. The primary differentiator may lie in subtle temporal features or pressure shifts related to gravitational resistance during uphill walking, which current spatial feature–focused models may not fully capture. In contrast, “slow walking” was classified with very high precision, indicating that variations in walking speed produce more distinct plantar pressure patterns than variations in surface uphill.

Figure 8. Confusion matrix for the encoder-augmented convolutional neural network (CNN).

Comparison With Prior Work

The use of deep learning, particularly CNNs, for plantar pressure analysis is consistent with recent trends in biomechanics and clinical research. Although many studies have successfully used CNNs for static pressure images (eg, for disease diagnosis), this research extends their application to dynamic gait classification. The accuracy achieved by the encoder-augmented CNN (96.21%) is competitive with or exceeds that reported in other studies using different sensor modalities or classification algorithms for similar tasks. The finding that an integrated AE-CNN architecture outperforms a standard CNN suggests that explicit feature learning and dimensionality reduction prior to classification can be a beneficial strategy for complex, high-variance data such as plantar pressure sequences. Clinical and translational implications should be interpreted with caution. This pilot study included only healthy young adults under controlled treadmill conditions and did not include participants with pathological gaits (eg, diabetes-related gait alteration). Therefore, while plantar pressure imaging is clinically relevant and the proposed framework shows technical promise, clinical screening performance has not been empirically validated here and requires future studies in clinical cohorts and real-world environments. In addition, our current approach performs frame-wise classification of stance-phase plantar pressure maps

and does not explicitly model temporal dynamics across gait cycles; future work could incorporate sequence models (eg, temporal CNNs or long short-term memory) to better capture temporal gait signatures.

Limitations

Although this study has yielded promising results, several limitations should be considered:

1. Single data source—plantar pressure data in this study were collected from a limited number of participants in a controlled laboratory environment. Whether the model's generalization ability can be extended to populations with different ages, genders, weights, or specific pathological gaits (eg, flat feet and diabetic foot) requires further validation.
2. Model interpretability—compared to classical machine learning models, deep learning models are often regarded as “black boxes,” with less transparent decision-making processes. Although this study validated the effectiveness of the model, it did not explore which specific regions or features of plantar pressure images the model used for classification.
3. Limited gait types—the study only covered 3 specific dynamic gaits. The applicability of the model to more complex daily activities, such as walking downhill, turning, or climbing stairs, has yet to be determined.

In addition, the evaluation in this study was reported primarily at the frame level. Although frame-wise metrics are useful for model comparison, participant-level performance (eg, aggregating frame-level predictions to participant-level decisions) should be reported in future studies to better reflect real-world use. Moreover, statistical uncertainty (eg, CIs via bootstrapping) and systematic robustness tests under controlled perturbations (eg, varying noise intensity or sensor shift) were not performed in this study and remain important directions for future work.

Conclusions

This study developed and evaluated a deep learning architecture named encoder-augmented CNN for gait classification using plantar pressure images. The model combines the feature extraction capabilities of an AE with the classification strengths of a CNN. Through systematic structural optimization, it ultimately achieved an accuracy of 96.21% in the 3-class dynamic-gait recognition task, outperforming classical machine learning methods and other deep learning variants in this study.

On the basis of the findings and limitations of this research, future studies could proceed in the following directions:

- Database expansion and model generalization—recruit a more diverse range of participants and collect data in

settings closer to real-life scenarios to validate and enhance the model's generalization ability. Future clinical studies could explore whether this framework can assist in the assessment of pathological gaits.

- Enhancing the interpretability of the model—introduce visualization techniques (eg, gradient-weighted class activation mapping) to analyze the plantar pressure heat maps that the model focuses on during classification decisions. This would improve understanding of the basis for the model's decisions and could potentially lead to the discovery of new biomechanical indicators.
- Multimodal data fusion—to address the confusion between fast walking and uphill walking, future work could attempt to fuse data from other sensors (such as gyroscopes and accelerometers) to build a multimodal gait recognition system, with the aim of achieving higher classification accuracy.
- Model lightweighting and real-time application—explore techniques such as model pruning or knowledge distillation to further reduce the computational complexity of the encoder-augmented CNN. This would enable its deployment on wearable devices or embedded systems for real-time gait monitoring and feedback.

Acknowledgments

No generative artificial intelligence tools were used in the writing or editing of this manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy and ethical restrictions but are available from the corresponding author on reasonable request. Access to deidentified data may be subject to institutional approval and a data use agreement, where applicable. Requests for data access may be directed to the corresponding author at byliau@fcu.edu.tw.

Funding

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

Conceptualization: CCC, BYL

Formal analysis: CCC

Investigation: CCC

Methodology: CCC, BYL, CWL, YKJ

Resources: CCC, BYL

Writing – original draft: CWL, YKJ, BYL

Writing – review & editing: CWL, YKJ, QQL, YYW, YSC, BYL

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CREMLS_checklist.

[\[DOCX File, 59 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Supporting optimization analyses, training curves, and hyperparameter tuning results.

[\[DOCX File , 554 KB-Multimedia Appendix 2\]](#)

References

1. Albuquerque P, Verlekar TT, Correia PL, Soares LD. A spatiotemporal deep learning approach for automatic pathological gait classification. *Sensors (Basel)*. Sep 16, 2021;21(18):6202. [FREE Full text] [doi: [10.3390/s21186202](https://doi.org/10.3390/s21186202)] [Medline: [34577408](https://pubmed.ncbi.nlm.nih.gov/34577408/)]
2. Bauman VV, Brandon SC. Gait phase detection in walking and stairs using machine learning. *J Biomech Eng*. Dec 01, 2022;144(12):121007. [doi: [10.1115/1.4055504](https://doi.org/10.1115/1.4055504)] [Medline: [36062965](https://pubmed.ncbi.nlm.nih.gov/36062965/)]
3. Krutaraniyom S, Sengchuai K, Booranawong A, Jaruenpun-yasak J. Pilot study on gait classification using machine learning. In: *Proceedings of the 2022 International Electrical Engineering Congress*. 2022. Presented at: iEECON; March 9-11, 2022; Khon Kaen, Thailand. [doi: [10.1109/ieecon53204.2022.9741586](https://doi.org/10.1109/ieecon53204.2022.9741586)]
4. Semwal VB, Jain R, Maheshwari P, Khatwani S. Gait reference trajectory generation at different walking speeds using LSTM and CNN. *Multimed Tools Appl*. Mar 13, 2023;82:33401-33419. [doi: [10.1007/s11042-023-14733-2](https://doi.org/10.1007/s11042-023-14733-2)]
5. Deschamps K, Matricali GA, Roosen P, Desloovere K, Bruyninckx H, Spaepen P, et al. Classification of forefoot plantar pressure distribution in persons with diabetes: a novel perspective for the mechanical management of diabetic foot? *PLoS One*. Nov 22, 2013;8(11):e79924. [FREE Full text] [doi: [10.1371/journal.pone.0079924](https://doi.org/10.1371/journal.pone.0079924)] [Medline: [24278219](https://pubmed.ncbi.nlm.nih.gov/24278219/)]
6. Amemiya A, Noguchi H, Oe M, Takehara K, Yamada A, Ohashi Y, et al. Relationship between elevated plantar pressure of toes and forefoot and gait features in diabetic patients. *Annu Int Conf IEEE Eng Med Biol Soc*. 2013;2013:4633-4636. [doi: [10.1109/EMBC.2013.6610580](https://doi.org/10.1109/EMBC.2013.6610580)] [Medline: [24110767](https://pubmed.ncbi.nlm.nih.gov/24110767/)]
7. Wang A, Li D, Fan N, Yuan S, Wu Q, Fu Z, et al. Piezoresistive-based gait monitoring technique for the recognition of knee osteoarthritis patients. *IEEE Access*. Nov 21, 2022;10:123874-123884. [doi: [10.1109/access.2022.3224047](https://doi.org/10.1109/access.2022.3224047)]
8. Nguyen ND, Bui DT, Truong PH, Jeong GM. Classification of five ambulatory activities regarding stair and incline walking using smart shoes. *IEEE Sens J*. May 17, 2018;18(13):5422-5428. [doi: [10.1109/jsen.2018.2837674](https://doi.org/10.1109/jsen.2018.2837674)]
9. Jeong GM, Truong PH, Choi SI. Classification of three types of walking activities regarding stairs using plantar pressure sensors. *IEEE Sens J*. Mar 15, 2017;17(9):2638-2639. [doi: [10.1109/JSEN.2017.2682322](https://doi.org/10.1109/JSEN.2017.2682322)]
10. Jun K, Lee S, Lee DW, Kim MS. Deep learning-based multimodal abnormal gait classification using a 3D skeleton and plantar foot pressure. *IEEE Access*. Nov 30, 2021;9:161576-161589. [doi: [10.1109/access.2021.3131613](https://doi.org/10.1109/access.2021.3131613)]
11. Johri S, Pratap S, Narayan J, Dwivedy SK. Sensory data classification for gait assessment using deep neural networks: a comparative study with SGD and Adam optimizer. In: *Proceedings of the 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation*. 2024. Presented at: IATMSI; March 14-16, 2024; Gwalior, India. [doi: [10.1109/iatmsi60426.2024.10503416](https://doi.org/10.1109/iatmsi60426.2024.10503416)]
12. Fricke C, Alizadeh J, Zakhary N, Woost TB, Bogdan M, Classen J. Evaluation of three machine learning algorithms for the automatic classification of EMG patterns in gait disorders. *Front Neurol*. May 21, 2021;12:666458. [FREE Full text] [doi: [10.3389/fneur.2021.666458](https://doi.org/10.3389/fneur.2021.666458)] [Medline: [34093413](https://pubmed.ncbi.nlm.nih.gov/34093413/)]
13. Cho H. Walking speed estimation and gait classification using plantar pressure and on-device deep learning. *IEEE Sens J*. Oct 1, 2023;23(19):23336-23347. [doi: [10.1109/jsen.2023.3305024](https://doi.org/10.1109/jsen.2023.3305024)]
14. Chhoeum V, Kim Y, Min SD. A convolution neural network approach to access knee joint angle using foot pressure mapping images: a preliminary investigation. *IEEE Sens J*. Aug 1, 2021;21(15):16937-16944. [doi: [10.1109/jsen.2021.3079516](https://doi.org/10.1109/jsen.2021.3079516)]
15. Ling ZQ, Zhang YP, Cao GZ, Chen JC, Li LL, Tan DP. AE-CNN-based multisource data fusion for gait motion step length estimation. *IEEE Sens J*. Nov 1, 2022;22(21):20805-20815. [doi: [10.1109/jsen.2022.3206883](https://doi.org/10.1109/jsen.2022.3206883)]
16. Ardhiyanto P, Subiakto RB, Lin CY, Jan YK, Liau BY, Tsai JY, et al. A deep learning method for foot progression angle detection in plantar pressure images. *Sensors (Basel)*. Apr 05, 2022;22(7):2786. [FREE Full text] [doi: [10.3390/s22072786](https://doi.org/10.3390/s22072786)] [Medline: [35408399](https://pubmed.ncbi.nlm.nih.gov/35408399/)]
17. Zhou H, Gui Y, Gu G, Ren H, Zhang W, Du Z, et al. A plantar pressure detection and gait analysis system based on flexible triboelectric pressure sensor array and deep learning. *Small*. Jan 2025;21(1):e2405064. [doi: [10.1002/sml.202405064](https://doi.org/10.1002/sml.202405064)] [Medline: [39473332](https://pubmed.ncbi.nlm.nih.gov/39473332/)]
18. F-Scan GO system. Tekscan. URL: <https://www.tekscan.com/products-solutions/systems/f-scan-system> [accessed 2026-02-14]

Abbreviations

AE: autoencoder

CNN: convolutional neural network

CREMLS: Consolidated Reporting of Machine Learning Studies

Edited by I Steenstra, A Mavragani; submitted 26.Nov.2025; peer-reviewed by MHF Aref, A Singhal; comments to author 08.Feb.2026; revised version received 31.Mar.2026; accepted 31.Mar.2026; published 21.Apr.2026

Please cite as:

Chang C-C, Lung C-W, Jan Y-K, Lu Q-Q, Wang Y-Y, Chen Y-S, Liao B-Y

Autoencoder-Enhanced Convolutional Neural Networks for Plantar Pressure-Based Gait Pattern Recognition: Model Development and Cross-Validated Evaluation Study

JMIR Form Res 2026;10:e88488

URL: <https://formative.jmir.org/2026/1/e88488>

doi: [10.2196/88488](https://doi.org/10.2196/88488)

PMID:

©Chuan-Chun Chang, Chi-Wen Lung, Yih-Kuen Jan, Qi-Qian Lu, Yi-You Wang, Yi-Sheng Chen, Ben-Yi Liao. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.