Learning and visualization of features using MC-DCNN for gait training considering physical individual differences

Yusuke OSAWA*, Keiichi WATANUKI*•**, Kazunori KAede•*** and Keiichi MURAMATSU*•**

*Graduate School of Science and Engineering, Saitama University.
255 Shimo-okubo, Sakura-ku, Saitama-shi, 338-8570 Japan
E-mail: y.osawa.949@ms.saitama-u.ac.jp
**Advanced Institute of Innovative Technology, Saitama University.
255 Shimo-okubo, Sakura-ku, Saitama-shi, 338-8570 Japan

Received: 11 July 2020; Revised: 17 September 2020; Accepted: 4 November 2020

Abstract

Several training methods have been developed to acquire motion information during real-time walking; these methods also feed the information back to the trainee. Trainees adjust their gait to ensure that the measured value approaches the target value, which may not always be suitable for each trainee. Therefore, we aim to develop a gait feedback training system that considers individual differences, classifies the gait of the trainee, and identifies adjustments for body parts and timing. A convolutional neural network (CNN) has a feature extraction function and is robust in terms of each feature position; therefore, it can be used to classify a gait as ideal or non-ideal. Additionally, when the gradient-weighted class activation mapping (Grad-CAM) is applied to the gait classification model, the output measures the influence degree contributed by the trainee’s each body part to the classification results. Thus, the trainee can visually determine the body parts that need to be adjusted through the use of the output. In this study, we focused on gaits related to stumbling. We measured the kinematics and kinetics data for participants and generated multivariate gait data, which were labeled as “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class using clustering with dynamic time warping. Next, the multichannel deep CNN (MC-DCNN) was used to learn the gait using the multivariate gait data and the corresponding classes. Finally, the data for verification were input into the MC-DCNN model, and we visualized the influence degrees of each place of the multivariate gait data for classification using Grad-CAM. The MC-DCNN model classified gaits with a high accuracy of 97.64±0.40%, and it learned the features that determine the thumb-to-ground distance. The output of the Grad-CAM indicated body parts, timing, and the relative strength of features that have an important effect on the thumb-to-ground distance.

Keywords: Healthcare, Gait training, Motion analysis, Walking factor, Neural network, Grad-CAM

1. Introduction

In Japan, which is a super-aged society, the number of people requiring long-term care is constantly increasing. Improving the quality of life of the elderly and extending their healthy life expectancy has become an urgent concern. In 2016, among the causes necessitating nursing care for the elderly aged 65 years or above, falls and fractures, and debilitation due to old age accounted for 12.5% and 13.8%, respectively; these two causes are considered the second highest after stroke and dementia (Japanese Ministry of Health, Labour and Welfare, 2016). Thus, fall prevention is a challenge even for active seniors. In the field of physical therapy, gait training using a walker, cane, and handrail is conducted under the guidance and supervision of a physical therapist; however, specialized knowledge is necessary to provide guidance. To overcome this problem, a gait training method that acquires information related to gait in real time and provides visual and auditory feedback to a trainee was proposed. Stanton et al. (2011) extracted 24 papers reporting 22 randomized trials on position, force, and electromyogram feedback for sitting, standing up, standing, and walking motions of a cerebral apoplexy patient, and they performed a systematic review with the meta-analysis. Further, they proved that feedback has a higher positive training effect than regular therapy or a placebo. In addition, they confirmed
that these training methods were more effective and sustainable over the long term compared to therapist-guided training methods. These studies targeted stroke patients, and therefore the gaits of the trainees are similar, and the goal of training is clear. For example, stroke patients are known to have smaller strides; thus, when the stride is targeted for training, the trainee may adjust the stride to be larger. However, target values are necessary when trainers of various gaits train changes related to the joint angle, driving force, and braking force via feedback training. However, similar to how physical therapists provide customized instructions to each trainee, the optimum values differ for each trainee because of physical individual differences (e.g., muscle strength and range of motion). Thus, it is necessary for a gait training method to present features of the ideal gait that needs to be satisfied by the trainee.

Our study aims at developing a gait feedback training system that considers individual differences. Thus, it is necessary to classify whether the gait of a trainee is ideal or non-ideal, and if it is non-ideal, to identify what part and timing of the gait should be adjusted, or how it should be adjusted. Machine learning is a method that learn patterns and features of data; there have been several studies that employ machine learning to classify human motions. For example, Martinez-Hernandez et al. (2018) input the measurements of wearable inertial sensors into a convolutional neural network (CNN) to detect a walking activity, and they predicted walking periods using a first-order Markov Chain. Lau et al. (2008) measured leg and foot movements during walking using kinematic sensors and classified the data into five walking conditions using a support vector machine (SVM): stair ascent, stair descent, flat ground, ascending slope, and descending slope. Further, Begg and Kamruzzaman (2005) measured the basic, kinetic, and kinematic gait data for young and old participants, and they classified the data into walking patterns of the young and old participants using the SVM. There are several studies that apply machine learning to human motion for detection and classification of motion. However, few studies have utilized the features learned by the machine learning models for motion training. Further, in the research on deep learning, methods have been developed to explain the basis of the classification when the deep learning model classifies input data, such as saliency maps (Ardizzone et al., 2013) and gradient-weighted class activation mapping (Grad-CAM) (Selvaraju et al., 2017). In previous studies, Grad-CAM is often used to explain where features learned by image classification models appear. Another study applied Grad-CAM to a time-series data classification model (Assaf and Schumann, 2019). However, even in these studies, the classification models were only used to explain the learned features. When Grad-CAM is applied to the data representing the gait, the output indicates the reasons for classifying the data as an ideal or non-ideal gait. Therefore, it is believed that the features of the ideal gait are not satisfied, and these features can be satisfied if the trainee makes adjustments to conform to the features. Thus, it may be possible to train each trainee to approach the ideal gait considering the physical individual differences among the trainers rather than adjusting to the absolute value as a goal.

In this paper, In this study, we evaluate the output of the presentation technique to represent the body part and timing that a trainee should adjust. The joint angle change and floor reaction force are measured when the participant walks freely or under restricted movement. Then, the data of one task is divided for each walking cycle, and this is converted into multivariate gait data. In addition, the thumb-to-ground distance is measured, and it is labeled as “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class using time-series clustering with dynamic time warping (DTW). A multichannel deep CNN (MC-DCNN) model learns the gait features using multivariate gait data as the input and the class corresponding to these data as the output. In addition, the gait of the verification data is classified using the MC-DCNN model, and the feature parts that are the basis of the classification are visualized using Grad-CAM.

Section 2 provides an overview of the developed gait feedback training system. Section 3 provides the measurement and generation of multivariate gait data used to construct a gait classification model, and Section 4 describes the learning of features of gait related to stumbling using MC-DCNN and the visualization of the part that is the basis of gait classification. Section 5 discusses the results, and Section 6 concludes.

2. Outline of gait feedback training system considering physical individual differences

The outline of a gait training system considering physical individual difference is shown in Fig. 1. First, the gait training system measures multivariate data representing the trainee’s gait. Next, the gait training system classifies the gait into ideal gait that does not require training or non-ideal gait that requires training, using a deep learning model based on the multivariate gait data of the trainee. The influence of each place of the multivariate gait data on the classification result is calculated. Then, the influence of each place is synthesized with the human body walking model, and it is fed
back to the trainee as visual information. The trainee intentionally adjusts the place presented visually. By repeating the above process, the trainee keeps adjusting so that the influence of the place with a large influence on the classification result becomes small. If the gait of the trainee is classified as a non-ideal gait, the place which has a large influence on the classification result shows the features of the non-ideal gait. Therefore, the trainee is expected to improve from the non-ideal gait to the ideal gait by walking so that the influence of the features become weak. Section 2.1 describes the generation of multivariate gait data, and Section 2.2 describes the definition of ideal and non-ideal gait in this study and the classification of gait. Section 2.3 describes the calculation of the influence of each place of multivariate gait data on the classification results.

2.1 Multivariate gait data generation
There is a considerable amount of information regarding gait. Muscle weakness with aging and body paralysis induce changes in the joint angle and the thumb-to-ground distance in terms of kinematics, the ground reaction force and joint moment in terms of dynamics, and muscle potential in terms of physiology. A small effect is recognized as the gait suitable for the individual; however, gait training is necessary when the effect becomes large, and it becomes a deviation motion when it exceeds the range permitted for the individual. The most frequently observed deviation motion has been identified by Perry and Burnfield (1993). According to their reports, excessive plantar dorsiflexion and eversion of the ankle joint; limited flexion and excessive flexion of the knee joint; excessive extension, excessive eversion, and excessive inversion; limited flexion and excessive flexion of the hip joint; excessive internal and external rotation; excessive adduction and abduction; excessive pelvic lifting; excessive retroversion and anteverision; excessive insufficient forward or posterior rotation and angular rotation; excessive forward and retroversion of the trunk; lateral flexion; and excessive forward or posterior rotation are observed. As shown by these deviation motions, trunk, hip joint, knee joint, and ankle joint angle are important as variables to determine gait. In this study, these joint angles were focused as variables to be trained from a kinematic perspective, and the ground reaction forces were focused from the dynamic viewpoint.

Fig. 2 shows the measurement of gait variables and generation of the multivariate gait data of the trainee. As shown in Fig. 2, the right direction of the trainee is defined as the X axis, the forward direction is defined as the Y axis, and the
vertical upward direction is defined as the Z axis. During gait training, the trainee walks freely on a treadmill with built-in force plates (Tec Gihan Co., Ltd., HPT -2200 D). Three components of the ground reaction forces are measured using two force plates built into the treadmill. The three-dimensional coordinates of the markers placed on the body of the trainee are measured using an optical three-dimensional motion analyzer (Natural Point Inc., OptiTrack), and then each joint (trunk, hip joint, knee joint, and ankle joint) angle is calculated from these marker coordinates. In the general gait analysis, the anatomical angles are calculated because the degree of freedom of the human joint is not always one. However, since the rotational motions of the body part are not measured in this study, it is impossible to express all the postures during walking at the anatomical angle. Therefore, in this study, the angles in the YZ and XZ planes in the local coordinate are calculated. And then, multivariate gait data are generated from these joint angles and ground reaction forces.

2.2 Classification into ideal and non-ideal gait considering physical individual differences

Perry and Burnfield (1993) measured the function of walking under the free walking condition in 420 healthy western men and women, and they determined the normal range (feature of averaging) for each gait variable in each age group. However, we do not need to walk on normal range if the gait is suitable to the physical individual difference (For example, muscle strength, range of motion, and length of each body part) and the environment. However, it is necessary to adjust the gait when it has the disadvantage by the motion which deviated too much. Typical examples of disadvantages caused by deviant movement include decreased stance stability, increased risk of stumbling and falling, decreased walking speed, decreased acceleration, and increased energy expenditure. It has been reported that many falls in the elderly are caused by stumbling while walking (Blake et al., 1988). Thus, it can be said that an increase in the risk of stumbling and falling is the main problem among the disadvantages caused by deviant movement. Therefore, in this study, ideal gait is defined as “gait rarely associated with stumbling”, and non-ideal gait is defined as “gait frequently associated with stumbling.”

This gait training system classifies a trainee's gait into “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class. Since stumbling involves the contact of toes with the ground or obstacles during the swing phase of walking, longer periods of maintaining a low thumb-to-ground distance are considered to be more likely to cause stumbling. Previous studies reported that the peak of the thumb-to-ground distance at the initial swing and the peak at the terminal swing decrease with age; further, they suggested that this increases the likelihood of stumbling in the elderly (Nishizawa et al., 1998). For this reason, the gait of the trainee can be classified in terms of stumbling by measuring the thumb-to-ground distance during walking. However, when gait classification is based only on the information of the thumb-to-ground distance, the trainee can forcibly raise the toe, for example bending the hip joint and knee joint excessively during the walking swing phase. Excessive toe rise can reduce the risk of stumbling, but it is not appropriate because it causes problems such as decreased stance stability. Therefore, it is desirable that the system classifies the gait related to the stumbling from the information of the multivariate gait data which is the compound data that shows the gait. Additionally, a representative value for each mean of group of gait rarely associated with stumbling and group of gait frequently associated with stumbling is used as a basis for the classification. It is necessary to appropriately adjust the basis value according to the height and the muscle force quantity if the physical individual differences of the trainees are to be considered. However, the number of parameters for determining the basis value for each trainee is enormous, and it is not realistic because measuring the length of each part of the body before the training is necessary to determine the precise unique value. In this study, the gait is classified using a machine learning model that extracts and learns the pattern and feature quantity of optional data and classifies the class of the data.

CNN is a machine learning method wherein a convolution layer and pooling layer are laminated. The convolution layer achieves feature extraction via processing, similar to spatial filtering. The pooling layer divides the input matrix into segments of the same size, and it outputs representative values in each segment. The pooling layer plays a role in ensuring robustness against translational movement of the representative value in the region of interest and in shortening the calculation time. The CNN demonstrates excellent performance in the field of image recognition; however, it is also used for time series data because of the abovementioned roles. Zheng et al. (2014) proposed MC-DCNN to learn features individually for each channel of the multichannel time series. MC-DCNN model learns the filter of the convolution layer for each channel of the multichannel time series and combines each output of the convolution layer for each channel in the fully connected layer. Multivariate gait data treated in this study can be regarded as multi-channel time series data in which gait variable corresponds to channel and each gait variable is a time series. From the characteristics of CNN, when learning gait by inputting multivariate gait data into MC-DCNN, the convolution layer corresponding to each gait variable
learns a filter for extracting feature quantities suitable for each gait variable, and the pooling layer obtains robustness for the time. In addition, the fully connected layer learns the weight of each value in the output matrix of the last convolution layer provided for each variable, that is, the relationship of gait variables and timing. Therefore, it is considered that by learning multivariate gait data using MC-DCNN, the gait classification model will perform classification considering individual differences. In this study, each multivariate gait data point for learning is labeled as “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class from corresponding thumb-to-ground distance, and the gait is learnt by inputting the multivariate gait data using MC-DCNN and outputting the class (Fig. 3). In gait training, the multivariate gait data of the trainee measured in real time are input into the MC-DCNN model, and the gait of the trainee is classified into “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class.

2.3 Visualization of body parts and timing to adjust

Methods to visualize the grounds of classification on the CNN include saliency maps and Grad-CAM. The Grad-CAM is a technique used to visualize the influence degree of each place in input on result of classification for each class as a heatmap. When arbitrary data is classified into any of several classes by using a classification model, an influence degree heat map, \( L^c \) corresponding to a specific class, \( c \), is obtained as follows

\[
L^c = \text{ReLU} \left( \sum_k \alpha_k A^k \right)
\]

\[
\alpha_k = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}
\]

\[
\text{ReLU}(x) = \max(x, 0)
\]

Fig. 3 Model structure of MC-DCNN for classify gait related to stumbling. This model learns the filter of the convolution layer for each channel of the multichannel time series, and combines each output of the convolution layer for each channel in the fully connected layer.

Fig. 4 Visualization of Influence degree to classification output by application of Grad-CAM to dog and cat classification model and the gait classification model, and Synthesis into a human walk model for intuitive visual feedback.
where \( y^c \) is the output of the output layer of class \( c \). \( A \) is the feature map that is the output of each filter of the convolution layer, and \( k \) is the number of filter of the target convolution layer. In general Grad-CAM, the last convolution layer is often the target. \((i, j)\) is the row and column number of the feature map, and \( Z \) is the product of \( i \) and \( j \). A weight coefficient \( a^c_{ij} \) on a \( k \)-th feature map of class \( c \) is calculated by Eq. (2), a feature map \( A^k \) multiplied by the weight coefficient \( a^c_{ij} \) is added, and an output by the activation function \( ReLU(\cdot) \) is defined as a heat map by Eq. (1). The location in the input matrix where the learning model is the basis of the classification can be understood using this heat map. Methods such as Grad-CAM, which visualize the influence of the CNN input layer or middle layer output on the CNN output, are often used to verify the features the constructed CNN model learns during image classification.

Figure 4 shows a heat map of a dog and cat input to a CNN model for discriminating between a dog and cat, the output provided by Grad-CAM, and an image of Grad-CAM applied to gait data. When the influence degree for classification result for dog are visualized on the image of the dog and cat, the head of the dog is emphasized. When this method is applied to multivariate gait data, the classification model classifies the gait considering the individual difference of the trainee, and the rows and columns of the heat map represent the influence degree of the timing and each variable during a walking cycle. So, it is considered that the trainee will visually understand the timing and body part to be adjusted by presenting them to the trainee. Since the MC-DCNN model in this study has a convolution layer for each gait variable, Grad-CAM calculates the influence degree of each variable to the output of MC-DCNN, and the output heat map is connected in the column direction. Further, because the output heat map is not intuitive, the body parts and timing that are active on the heat map are synthesized to the human body model, and they are presented on the display arranged in front of the treadmill.

3. Experiments for generating multivariate gait data

3.1 Measurement of each gait variable during walking

This experiment was approved by Saitama University Ethics Committee (Approval Number: H29-E-12). For safety reasons, eight “healthy” Japanese men (mean age 23.9 ± 1.0 years old) participated in the study. The participants provided informed consent. The infrared cameras of a three-dimensional motion analyzer were arranged around a treadmill with a built-in force plate. Two force plates are built under the right and left belts of the treadmill. To increase the types of gait, the experimental conditions were set as (1) “Normal walking” and (2) “Restricted walking (walking with muscle load and limited joint motion)” assuming the gait of the elderly. The reflection markers were placed on the body of the participants as shown in Fig. 5, and under condition (2), participants wore weights on the anterior part of the upper extremities, wrists, and the ankle joints, and braces that limited joint motion on the elbow and knee joints (Sanwa Manufacturing Co., Ltd, Expert set III of teaching materials for simulating elderly) as shown in Fig. 6. Under each condition, participants walked on the treadmill for 120 s after following a reference posture for 5 s. The reference posture was defined as a posture with the back straight, legs shoulder-width apart, and the hands kept away from the body at a width of a closed fist. Under each condition, the speed of the treadmill was automatically controlled so that the anteroposterior coordinates of the participant remained in the center on the treadmill to allow them to walk at a natural walking speed. The three-dimensional coordinates (X, Y, Z) of each marker and the three components of the ground
reaction forces (X, Y, Z) were measured using a three-dimensional motion analyzer and the force plates, at each sampling frequency of 250 Hz. One trial was conducted for each condition.

3.2 Preprocessing for generation of input data used for learning

Because there were some moments when the three-dimensional coordinates could not be measured because of the effect of external light or because the reflection markers were blocked by the bodies of the participants, spline interpolation was performed on the three-dimensional coordinates of the markers. As shown in Fig. 7, each vector forming each joint angle was obtained from the three-dimensional coordinates of each marker, and the joint angles (Trunk, hip, knee, and ankle) in the YZ and XZ planes were calculated in the range of 0 to 360 degrees using the inner product and the outer product between the vectors. The Z coordinates of the markers placed at the first metatarsal heads of the right and left foot were defined as the thumb-to-ground distances. Since the calculated angle and thumb-to-ground distance are affected by the difference in the position of the markers on the body, we calculated these changes from the reference posture using the difference between the values obtained during walking and the reference posture.

Next, the three components of the ground reaction force were normalized by the body weight of each participant in condition 1 data, and by the sum of the body weight of each participant and the weight of the orthosis in condition 2 data. In this study, the moment when the Z component of the normalized ground reaction force becomes greater than 0.05 was detected as the right foot contact. To generate input data for learning, each variable change during the 120 s was divided into data for each stride from the initial right foot contact to the next right foot contact. The matrix, which was obtained by connecting these variable changes (trunk joint angle in the YZ and XZ planes, right and left hip joint angles in the YZ and XZ planes, right and left knee joint angles in the YZ and XZ planes, and right and left ankle joint angles in the YZ and XZ planes, and right and left ground reaction forces 3 components (X, Y, Z)) in the row direction was used as multivariate gait data. Further, the left foot contact was detected in the same manner, and the data were divided accordingly. To treat the multivariable gait data divided by the left foot contact and the multivariable gait data divided by the right foot contact in the same manner, the left and right rows of the data divided by the left foot contact were replaced, and the signs of the angle in the XZ plane were switched. Thus, multivariate gait data points were generated for 1,548 strides. Each variable in the generated data was normalized with the maximum and minimum values of each variable in all the data. In addition, because the length in the column direction was different for the multivariate gait data divided by stride, each data was normalized such that the walking cycle became 100%. Finally, a lowpass filter was applied at a cutoff frequency of 30 Hz.

4. Learning the features of “gait rarely associated with stumbling” and “gait frequently associated with stumbling”

4.1 Labeling multivariate gait data as “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class using time-series clustering with dynamic time warping (DTW)

When learning a classification problem using CNN, a correct answer class corresponding to input data is required as an output. In this study, we input multivariate gait data of gait trainees and classify them into “gait rarely associated with stumbling” class or “gait frequently associated with stumbling” class. Therefore, each input data point used for learning must be labeled to these gait.
In multivariate gait data measurement, two conditions—normal walking and restricted walking (walking with muscle load and limited joint motion)—were considered; however, the gait in normal walking is not always the “gait rarely associated with stumbling” class, and the gait in restricted walking is not always the “gait frequently associated with stumbling” class, and it depended on the muscular strength and mood of the participants in the experiment. Therefore, multivariate gait data were labeled based on the corresponding thumb-to-ground distance. In addition, even in a single trial with the same participants and same conditions, the gait of each stride was not always the same because of the effect of fatigue, and therefore the toe-to-ground distance was clustered using unsupervised learning, to label the multivariate gait data used for learning.

The k-shape method using shape-based distance (SBD) (Paparrizos and Gravano, 2016) as the distance scale and the nearest neighbor method using DTW as the distance scale (Sakoe and Chiba, 1978) are examples of clustering methods employed for time-series data. The SBD method uses standardized cross-correlation, and the k-shape method is applied to the clustering of time series data to be considered for scaling and phase shift. The DTW distance is the sum of the distances at which the sum of the distances of the warping paths connecting both ends of the distance matrix is minimum when the distance matrix is created based on the distance in the combination of all points of the two-time series. The DTW distance is applied to the clustering of time series data having different lengths and time series data having a phase shift. The k-shape method, which considers scaling, is not suitable for clustering the thumb-to-ground distance because the absolute value of the thumb-to-ground distance affects the ease of stumbling.

In this study, we performed clustering of the thumb-to-ground distance using the nearest neighbor method using DTW. The number of clusters was determined by the elbow method (Bholowalia and Kumar, 2014). The elbow method is one of the methods to determine the optimum number of clusters in the clustering method and calculates the sum of squared errors (SSE) representing the distance between the center of gravity of each cluster and the data classified into the clusters when the arbitrary number of clusters is set. When the number of clusters increases sufficiently, SSE decreases and settles to a constant value. Therefore, in the Elbow method, the number of clusters in which the decrease in SSE becomes extremely small when the number of clusters is increased is determined as the optimum number of clusters. The result of the elbow method is shown in Fig. 8. The number of clusters of the thumb-to-ground distance is two in which the decrease of SSE becomes extremely small. Fig. 9 shows all the thumb-to-ground distances classified into each cluster, as well as their mean and standard deviation. Cluster 0 contained 790 data points and Cluster 1 contained 758 data points. The thumb-to-ground distance of Cluster 0 is larger than that of Cluster 1, and the difference is especially large in the terminal swing. Therefore, multivariate gait data corresponding to the thumb-to-ground distance classified in Cluster 0 were labeled as the “gait rarely associated with stumbling” class, and those corresponding to the thumb-to-ground distance classified in Cluster 1 were labeled as “gait frequently associated with stumbling” class.

![Fig. 8 Determination of the number of clusters of the thumb-to-ground distance by results of the elbow method.](image1)

![Fig. 9 All the thumb-to-ground distance classified into each cluster (grey) and their mean and standard deviation (red or blue).](image2)

[DOI: 10.1299/jbse.20-00337]
Fig. 10 All the data of each variable in each cluster (grey) and their mean and standard deviation (red or blue).
X axis is walking cycle [%], and y axis is each preprocessed gait variable (left two columns: trunk angle in the YZ plane, trunk angle in the XZ plane, right hip joint angle in the YZ plane, right hip joint angle in the XZ plane, left hip joint angle in the YZ plane, left hip joint angle in the XZ plane, right knee joint angle in the YZ plane, right knee joint angle in the XZ plane, left knee joint angle in the YZ plane, left knee joint angle in the XZ plane, right ground reaction force X component, right ground reaction force Y component, right ground reaction force Z component, left ground reaction force X component, left ground reaction force Y component, left ground reaction force Z component).
Figure 10 shows the values of each variable of the multivariate gait data corresponding to the thumb-to-ground distance classified into each cluster, and their mean and standard deviation. When Cluster 0 and Cluster 1 are compared, there is a difference in the following variables. Cluster 1 is smaller in the total walk cycle of the trunk angle in the YZ plane, after 60% of the right knee joint angle in the YZ plane, before 40% of the left knee joint angle in the YZ plane, and in the range of 50–80% of the right ankle joint angle in the YZ plane, 0–30% of the left ankle joint angle in the YZ plane, 0–10% and 90–100% of the absolute right ankle joint angle in XZ plane, and 40–60% of the absolute left ankle joint angle in the XZ plane. Cluster 1 is smaller in the absolute values of the first and second peaks of the ground reaction force Y component in both legs, and larger in the absolute value of the local minimum of the ground reaction force Z component in both legs. Further, even in the variable in which a remarkable difference is not observed in the mean value, the change in the quantity of the angle in the YZ plane tends to be smaller for each joint in Cluster 1; the change in the quantity of the angle in the XZ plane tends to be larger.

4.2 Learning multivariate gait data using MC-DCNN

Multivariate gait data were input, the clustering results of the thumb-to-ground distance corresponding to each multivariate gait data point were output, and the features of “gait rarely associated with stumbling” and “gait frequently associated with stumbling” were learned using MC-DCNN. The model structure of the MC-DCNN is shown in Fig. 11 and summarized in Table 1. The number of epochs was 3,000, the optimization function was stochastic gradient descent, and the error function was categorical cross entropy. For the verification of the influence degree visualization, 10 data points were randomly selected for each class from a total of 1,548 data points, and 80% of the 1,528 data points excluding the 20 data points were used as the training data; 20% were used as validation data. The model with the lowest validation loss among 3,000 epochs was saved using the model check point function. The learning was performed 10 times, and the average accuracy of gait classification related to stumbling for validation data in 10 gait classification models generated was $97.64 \pm 0.40\%$.

![Diagram of MC-DCNN model](image_url)

**Table 1** Layer description of the MC-DCNN model for gait classification.

| Layer name       | Layer description                                                  |
|------------------|--------------------------------------------------------------------|
| Input            | $20 \times 400 \times 1$ Two-dimensional data                     |
| Conv1d n_1       | Filters: 32, kernel size: 25, stride: 1, padding: causal, activation: relu |
| Max pooling1d n_1| pool size: 2, padding: valid                                      |
| Conv1d n_2       | Filters: 64, kernel size: 10, stride: 1, padding: causal, activation: relu |
| Max pooling1d n_2| pool size: 2, padding: valid                                      |
| Dropout          | Rate: 0.25                                                         |
| Flatten          |                                                                    |
| Dense 1          | Units: 256, activation: relu                                       |
| Dense 2          | Units: 64, activation: relu                                       |
| Dense 3          | Units: 16, activation: relu                                       |
| Output           | Units: 2, activation: softmax                                     |
4.3. Visualization of the influence degree for the classification results obtained using Grad-CAM

For the model with the lowest validation loss of 10 learning models, 20 data points were input for the verification of the influence degree visualization, and the gait classes were classified. In addition, Grad-CAM was used to visualize the location of the features that constitute the basis for the classification.

The classification result and output of the softmax function of the output layer when the data for feature visualization verification of each was input to the MC-DCNN model are summarized in Table 2. The classification of gait related to stumbling resulted in the correct classification in all 20 data points. In particular, the output of softmax for the class of data point 3 and data point 16 are the largest in each class, and those of data point 4 and data point 17 are intermediate among all data. Since the softmax function outputs the ratio of the elements of the input vector, the output of the softmax function indicates the ratio of the feature strength of each class of the input data. Figure 12 shows the right thumb-to-ground distance and the value of each variable for these four data, and Fig. 13 shows the Grad-CAM output that is the influence degree on the output of the softmax function which is the output layer for each class of each place in these data. These heat maps are normalized by the maximum value of the influence degree of each data. This MC-DCNN model focuses on the trunk angle in the YZ plane, left hip joint angle in the YZ plane, right knee joint angle in the XZ plane, and left ground reaction force X component as grounds for classification of gait rarely associated with stumbling. In addition, it focuses on the trunk angle in the XZ plane, right knee joint angle in the YZ plane, left knee joint angle in the YZ plane, left knee joint angle in the XZ plane, right ankle joint angle in the YZ plane, left ankle joint angle in the XZ plane, right ground reaction force Y component, right ground reaction force Z component, and left ground reaction force Y component as grounds for classification of gait frequently associated with stumbling. In data point 3, the trunk

Table 2 Classification result of multivariate gait data and output of softmax function using MC-DCNN model for gait classification.

| Class of input data                  | Data point No. | Result of classification | Output of soft max for the “gait rarely associated with stumbling” (×10⁻¹) | Output of soft max for the “gait frequently associated with stumbling” (×10⁻¹) |
|-------------------------------------|----------------|----------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
|                                     | 1              | Correct                    | 9.9987                                                                          | 0.0013                                                                          |
| “Gait rarely associated with stumbling” | 2              | Correct                    | 9.9984                                                                          | 0.0016                                                                          |
|                                     | 3              | Correct                    | 9.9996                                                                          | 0.0004                                                                          |
|                                     | 4              | Correct                    | 5.3405                                                                          | 4.6595                                                                          |
|                                     | 5              | Correct                    | 7.5767                                                                          | 0.4233                                                                          |
|                                     | 6              | Correct                    | 9.9992                                                                          | 0.0008                                                                          |
|                                     | 7              | Correct                    | 9.9987                                                                          | 0.0013                                                                          |
|                                     | 8              | Correct                    | 9.8296                                                                          | 0.1704                                                                          |
|                                     | 9              | Correct                    | 9.9993                                                                          | 0.0007                                                                          |
|                                     | 10             | Correct                    | 8.1583                                                                          | 1.8417                                                                          |
| “Gait frequently associated with stumbling” | 11             | Correct                    | 0.0017                                                                          | 9.9983                                                                          |
|                                     | 12             | Correct                    | 0.0001                                                                          | 9.9999                                                                          |
|                                     | 13             | Correct                    | 3.9809                                                                          | 6.0191                                                                          |
|                                     | 14             | Correct                    | 0.0026                                                                          | 9.9974                                                                          |
|                                     | 15             | Correct                    | 0.0035                                                                          | 9.9965                                                                          |
|                                     | 16             | Correct                    | 0.0001                                                                          | 9.9999                                                                          |
|                                     | 17             | Correct                    | 4.6168                                                                          | 5.3832                                                                          |
|                                     | 18             | Correct                    | 0.0003                                                                          | 9.9997                                                                          |
|                                     | 19             | Correct                    | 0.6475                                                                          | 9.3525                                                                          |
|                                     | 20             | Correct                    | 0.0001                                                                          | 9.9999                                                                          |
angle in the YZ plane angle strongly influenced the gait rarely associated with stumbling, and the first half of the left knee joint angle in the YZ plane and the second half of the right ankle joint angle in the YZ plane moderately influenced the gait frequently associated with stumbling. In data point 16, the trunk angle in the YZ plane moderately influenced the gait rarely associated with stumbling, the knee joint angle in the YZ plane moderately influenced the gait frequently associated with stumbling, and the right ankle joint angle in the YZ plane strongly influenced the gait frequently associated with stumbling. In data point 4, the trunk joint angle in the YZ plane moderately influenced the gait rarely associated with stumbling, the first half of the left knee joint angle in the YZ plane moderately influenced the gait frequently associated with stumbling, and the right ankle joint angle in the YZ plane strongly influenced the gait frequently associated with stumbling. In data point 17, the trunk joint angle in the YZ plane moderately influenced the gait rarely associated with stumbling, the right knee joint angle in the YZ plane strongly influenced the gait frequently associated with stumbling, and the left knee joint angle in the YZ plane and the right ankle joint angle in the YZ plane moderately influenced the gait frequently associated with stumbling.

Fig. 12 Preprocessed Thumb-to-ground distance and each preprocessed variate data for the verification of the influence degree visualization (red is data point 3, orange is data point 4, blue is data point 16, and green is data point 17) (column 1: trunk angle in the YZ plane, trunk angle in the XZ plane, right hip joint angle in the YZ plane, right hip joint angle in the XZ plane, left hip joint angle in the YZ plane) (column 2: left hip joint angle in the XZ plane, right knee joint angle in the YZ plane, right knee joint angle in the XZ plane, left knee joint angle in the XZ plane, left knee joint angle in the YZ plane) (column 3: right ankle joint angle in the YZ plane, right ankle joint angle in the XZ plane, left ankle joint angle in the YZ plane, left ankle joint angle in the XZ plane, right ground reaction force X component) (column 4: right ground reaction force Y component, right reaction force Z component, left ground reaction force X component, left ground reaction force Y component, left ground reaction force Z component).
Fig. 13 Influence degree heatmaps on the output of the softmax function, which is the output layer, for each class of each place in the input multivariate gait data (GRF means ground reaction force). These influence degrees were normalized by the maximum value of each data.
5. Discussion

5.1. Comparison between “gait rarely associated with stumbling” class and “gait frequently associated with stumbling” class labeled using DTW

In this study, the thumb-to-ground distances were clustered by the nearest neighbor method using DTW, and the corresponding multivariate gait data were labeled as the “gait rarely associated with stumbling” class or the “gait frequently associated with stumbling” class. We compare the mean and standard deviation before normalization of the thumb-to-ground distance in each class. In the “gait rarely associated with stumbling” class, the peak in the initial swing was 22.1 ± 11.2 mm, the local minimum in the mid swing was 10.2 ± 8.5 mm, and the peak in the terminal swing was 100.0 ± 10.7 mm. Whereas, in the “gait frequently associated with stumbling” class, the peak in the initial swing was 13.0 ± 9.0 mm, the local minimum in the mid swing was 9.1 ± 8.0 mm, and the peak in the terminal swing was 65.2 ± 13.3 mm. In our previous experiment (Osawa et al., 2017) that compares the thumb-to-ground distance of young and elderly people, the peak in the initial swing was 26.9 ± 6.8 mm, the local minimum in the mid swing was 13.1 ± 4.8 mm, and the peak in the terminal swing was 99.9 ± 7.6 mm in the young group. On the other hand, in the elderly people, the peak in the initial swing was 26.1 ± 9.5 mm, the local minimum in the mid swing was 15.5 ± 7.3 mm, and the peak in the terminal swing was 88.0 ± 20.1 mm. The thumb-to-ground distance of the “gait rarely associated with stumbling” class was similar to the young group, and the thumb-to-ground distance of the “gait frequently associated with stumbling” class was lower than the elderly group. It is impossible to express the degree of stumbling because the risk of stumbling is affected not only by the thumb-to-ground distance but also by the environment of the ground and the cognitive ability to notice obstacles. However, from the viewpoint of the thumb-to-ground distance, the multivariate gait data included in the “gait frequently associated with stumbling” class has the risk of stumbling, which is similar to that of elderly people.

Comparing the mean of each variable of the multivariate gait data included in each class from Fig. 10, the trunk angle in the YZ plane of the “gait frequently associated with stumbling” class is small in the entire walking cycle, and the gait of this class is the anteverision posture. And the right and left knee joint angle in the YZ plane in the mid swing is small, and it is large in the terminal swing. In addition, the right and left ankle joint angle in the YZ plane in the pre-swing is small, and the absolute value of ground reaction force Y component in the terminal stance is small. This fact means that the gait of the “gait frequently associated with stumbling” class has insufficient plantar flexion in the pre-swing, a small driving force, the insufficient knee joint flexion in the mid swing, and the insufficient knee joint extension in terminal swing. There is the data that the absolute value of right and left ankle joint angle in the XZ plane in the terminal swing is small and that the ankle joint angle in the XZ plane is inverted in the whole walking cycle. This indicates that the participant abducts the ankle joint in the “gait frequently associated with stumbling” class. Although there is no large difference in the mean value, the standard deviation is large for each variable of the “gait frequently associated with stumbling” class (For example, the trunk angle in the XZ plane, the hip joint angle in the XZ plane, the knee joint angle in the XZ plane, the ground reaction force X component). This means that, in the gait of the “gait frequently associated with stumbling” class the trunk is swung in the lateral direction, and the hip joint is abducted or there is an external rotation.

5.2. Comparison of the data for the verification of influence degree visualization and the influence degree of each place of multivariate gait data using Grad-CAM on the output

The gait classification model constructed in this experiment had a high classification accuracy of 97.64 ± 0.40. When comparing the thumb-to-ground distance of the four data input into Grad-CAM, data point 1 with the highest score as a “gait rarely associated with stumbling” class is the highest in the terminal swing, and data point 16 with the highest score as a “gait frequently associated with stumbling” class is the lowest. Further, data points 4 and 17 with the intermediate scores on each class are intermediate values. Despite the fact that learning was conducted using multivariate gait data with binary values of 0 and 1 as output, the score was output as intermediate for the data in which the thumb-to-ground distance was in the middle. This implies that the MC-DCNN model learned the features to determine the thumb-to-ground distance, and it can be said that the level of the stumbling of the gait of the trainee is determined from the score.

Further, When comparing the thumb-to-ground distance of the four data input into Grad-CAM, the data point 16 with the highest output score as the “gait frequently associated with stumbling” class had, in particular, small flexion of the knee joint angle in the YZ plane in the mid swing leg, small braking force and driving force of the ground reaction force Y component, large change of the trunk joint angle in the XZ plane, the large hip joint angle in the XZ plane, large change
of the knee joint angle in the XZ plane and inverted ankle joint angle in the XZ plane for other data, and large ground reaction force X component. This indicates that the gait of data point 16 has insufficient flexion of the knee joint in mid-swing, abduction and external rotation of the hip joint and abduction of the ankle joint, and it shows “the circumduction gait.” When the mean of the “gait rarely associated with stumbling” class is presented to the trainee as a target value in the training, it is necessary for the participant of the data point 16 to adjust knee joint flexion angle, driving force of ground reaction force, and ankle joint angle in the XZ plane in the mid swing because the difference described in the Section 5.1 exists between the classes. However, in the heat map of the influence degree on the output score of each class by Grad-CAM, it is found that the knee joint angle in the YZ plane and the ground reaction force Y component affect the output score, but the influence degree on the ankle joint angle in the XZ plane is not observed. In addition, there is an influence on the output score in the trunk angle in the XZ plane, which was not large different between each mean of the classes, while there is no influence in the hip joint angle in the XZ plane and the knee joint angle in the XZ plane, which were not large different between each mean of the classes. These results indicate that the gait classification model for stumbling learned not only the mean of the classes but also the features that determine the thumb-to-ground distance from the shape of the waveform and the relationship among the variables. When attention is paid to “stumbling” as a disadvantage, the model judged that the abduction of the ankle joint, the abduction and external rotation of the hip joint, which are represented by the ankle joint angle in the XZ plane, hip joint angle in the XZ plane, and knee joint angle in the XZ plane are acceptable movements, and that the gait is suitable for individuals. The gait training system shows the priority of the position to be adjusted for the trainee because the influence degree on the output score is different based on each body part or timing in the influence degree heat map of each multivariate gait data. The influence degree heat map shown in Fig. 13 is not intuitive for the trainee; therefore, it is necessary to synthesize it into a body walking model. By using the body walking model to demonstrate this effect to the trainee, this gait training system can show the position which affects the stumbling and the priority of the adjustment, considering the gait allowed by the individual, and it can provide guidance that is similar to the guidance of a physical therapist. This system enables the gait trainee to understand the body part and timing to be adjusted by him/herself. Therefore, it can be said that the gait trainee can efficiently perform gait training for stumbling compared to the case in which only the thumb-to-ground distance is presented, or the mean of each gait variable is presented.

5.3. Limitations

Comparing the output heat maps of the Grad-CAM, the normalized influence degree of the right knee joint angle in the YZ plane of data point 17, which is the intermediate gait, is larger than that of data point 16, which is the most frequently associated with stumbling, whereas the knee joint angle in the YZ plane of the data point 16 is especially small in the mid swing. This is because the entire output is normalized by the maximum value of the output of each iteration of Grad-CAM when the Grad-CAM is heat mapped. The maximum value of the influence degree of each data was as follows. It for data point 3 was $1.13 \times 10^{-8}$, it for data point 4 was $6.55 \times 10^{-4}$, it for data point 16 was $2.42 \times 10^{-12}$, and it for data point 17 was $6.13 \times 10^{-4}$. This means that the influence degree of each location is small in the data with a large output score for each class, and the influence degree of each location is large in the data with an intermediate output score. In classification problems, such as the MC-DCNN model in this study, the softmax and sigmoid functions are often used for the activation function of the output layer. The output of the softmax function is given by

$$y_k = \frac{\exp(a_k)}{\sum_{i=1}^{n} \exp(a_i)} \tag{4}$$

where $a$ is an input vector. As shown in Fig. 14, the softmax function is a more general version of the sigmoid function, where any vector input is normalized to fit the output into the range 0 to 1. The Grad-CAM calculates the gradient of a given convolution layer with respect to the output of the output layer. The gradient becomes gentle as the input becomes smaller or larger, and the slope is the largest when it is in the middle. Therefore, the non-normalized output of Grad-CAM for data points 4 and 17, which are intermediate scores, is large, and that for data points 3 and 16, which are the highest scores, are small. Further, even if the identity function—often used in regression problems—is used instead of the softmax and sigmoid functions, the output of Grad-CAM only indicates the contribution to the output score, and it does not suggest the absolute strength of the feature of each place of the input data. Thus, it is not possible to compare the absolute strength of each feature among different data in the output of Grad-CAM. The Grad-CAM represents the relative strength of a feature in each input data by normalizing it with the maximum value of each output. In addition,
since the output of the Grad-CAM is the influence degree of the output of the specified intermediate layer on the output score of the learning layer, it can indicate the body part and timing to be adjusted by the trainee, but it does not show how to adjust it. Therefore, it is necessary for the trainee to adjust by trial and error and to confirm the classification result.

In this experiment, we measured the multivariate gait data of young people during normal walking and restricted walking. Therefore, the developed model to classify gait is not necessarily applicable to the gait of elderly people. In actual training, it is necessary to use a model constructed by measuring the gait of the elderly to classify gait. In addition, it is necessary to examine other symptoms (e.g., hemiplegic gait) because the objects of the classification are gaits rarely and frequently associated with stumbling. In this experiment, the classification model was constructed using the multivariate gait data of only man. Though the number of participants in the experiment artificially increased by carrying out two conditions, the total number of participants was less, with eight persons. Since gait is known to be affected by age, gender, etc., it is necessary to measure multivariate gait data of various experiment participants, and to learn features on stumbling and to construct a gait classification model to develop a more general gait training system. The classification model learned from the multivariate gait data of various humans seems to make the trainee train the gait on the stumbling considering various individual differences.

6. Conclusion

The MC-DCNN model for classifying gait was able to classify the gait related to stumbling with high accuracy (97.64 ± 0.40%), and the data for verification input into the model and the places that were grounds for classification of these data were visualized using Grad-CAM. The gait classification model learned the features to determine the thumb-to-ground distance, and the output of the Grad-CAM was able to visualize the places that affect the classification result of the “gait rarely associated with stumbling” class or the “gait frequently associated with stumbling” class in the multivariate gait data and their relative strengths. These results indicate that the proposed method is more similar to the guidance of physical therapists than the conventional method in which only one gait variable is presented, or the method in which the mean of the gait variables is set as a target value, by visualizing the body part and timing which cause stumbling. It also suggests that the trainee can efficiently undergo walking training considering the physical individual difference by adjusting these parts and timing.

The method proposed in this paper visualizes the body part and timing to be adjusted by the trainee, but it does not provide a quantitative value to the amount of adjustment required. In the future, we aim to develop a method that sets the target value considering the individual difference. We will measure multivariate gait data of the elderly and construct a model to classify gait from the data. The training effects of the proposed method and the gait training method using general motion information feedback will be compared. Further, we aim to develop a training system for various abnormal gaits (e.g., hemiplegic gait) by constructing a model for symptoms other than stumbling. It is expected that the proposed method can be used not only in the field of nursing but also in the field of sports, if the model is trained appropriately.

References

Ardizzone, E., Bruno, A. and Mazzola, G., Saliency based image cropping, International Conference on Image Analysis and Processing (ICIAP2013) (2013), pp. 773–782, DOI: https://doi.org/10.1007/978-3-642-41181-6_78.
Assaf, R. and Schumann, A., Explainable deep neural networks for multivariate time series predictions, Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (2019), pp. 6488–6490, DOI: https://doi.org/10.24963/ijcai.2019/932.

Begg, R. and Kamruzzaman, J., A machine learning approach for automated recognition of movement patterns using basic kinetic and kinematic gait data, Journal of Biomechanics, Vol. 38 (2005), pp.401–408, DOI: https://doi.org/10.1016/j.jbiomech.2004.05.002.

Bholowalia, P. and Kumar, A., EBK-means: A clustering technique based on elbow method and k-means in WSN, International Journal of Computer Applications, Vol. 105, No. 9 (2014), pp. 17–24, DOI: https://doi.org/10.5120/18405-9674

Blake, A. J., Morgan, K., Bendall, M. J., Dallosso, H., Ebrahim, S. B. J., Arie, T. H. D, Fentem, P. H. and Bassey, E. J., Falls by elderly people at home: Prevalence and associated factors, Age and Ageing, Vol. 17 (1988), pp. 365–372, DOI: https://doi.org/10.1093/ageing/17.6.365.

Japanese Ministry of Health, Labour and Welfare, Summary Report of Comprehensive Survey of Living Conditions 2016 (2016).

Lau, H., Tong, K. and Zhu, H., Support vector machine for classification of walking conditions using miniature kinematic sensors, Medical & Biological Engineering & Computing, Vol. 46 (2008), pp. 563–573, DOI: https://doi.org/10.1007/s11517-008-0327-x.

Martinez-Hernandez, U., Rubio-Soils, A. and Dehghani-Sanji, A. A., Recognition of walking activity and prediction of gait periods with a CNN and first-order MC strategy, 2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob) (2018), pp. 897–902, DOI: https://doi.org/10.1109/biorob.2018.8487220.

Nishizawa, S., Nagasaki, H., Furuna, T., Okuzumi, H., Sugiura, M., Ito, H. and Fujita, M., Foot clearance during walking of older adults in the community, Biomechanisms, Society of Biomechanisms Japan, Vol. 14 (1998), pp. 69–79, DOI: https://doi.org/10.3951/biomechanisms.14.69 (in Japanese).

Osawa, Y., Watanuki, K., Kaede, K., Factorial analysis of stumbling using evaluation of lower limb trajectory during walking, Proceeding of Mechanical Engineering Congress (2017), The Japan Society of Mechanical Engineering, DOI: https://doi.org/10.1299/jsmemecj.2017.S1150102

Paparrizos, J. and Gravano, L., k-shape: Efficient and accurate clustering of time series, Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (2015), pp. 1855–1870, DOI: https://doi.org/10.1145/2723372.2737793.

Perry, J. and Burnfield, J.M., Gait analysis: normal and pathological function, Developmental Medicine and Child Neurology, Vol. 35 (1993) pp. 1122–1122.

Sakoe, H. and Chiba, S., Dynamic programming algorithm optimization for spoken word recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 26 (1978), pp. 43–49, DOI: https://doi.org/10.1109/tassp.1978.1163055.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D., Grad-CAM: Visual explanations from deep networks via gradient-based localization, The IEEE International Conference on Computer Vision (ICCV) (2017), pp. 618–626, DOI: https://doi.org/10.1109/iccv.2017.74.

Stanton, R., Ada, L., Dean, C. M. and Preston, E., Biofeedback improves activities of the lower limb after stroke: A systematic review, Journal of Physiotherapy Vol. 57 (2011), pp. 145–155, DOI: https://doi.org/10.1016/s1836-9553(11)70035-2/.

Zheng, Y., Liu, Q., Chen, E., Ge, Y. and Zhao, J. L., Time series classification using multi-channels deep convolutional neural networks, International Conference on Web-Age Information Management (2014), pp. 298–310, DOI: https://doi.org/10.1007/978-3-319-08010-9_33.