


Article

A Deep-Learning Approach for Identifying a Drunk Person Using Gait Recognition

Suah Park ¹, Byunghoon Bae ¹, Kyungmin Kang ¹, Hyunjee Kim ¹, Mi Song Nam ¹, Jumyung Um ^{2,*} and Yun Jung Heo ^{1,3,*}

¹ Department of Mechanical Engineering, College of Engineering, Kyung Hee University, 1732 Deokyoung-daero, Giheung-gu, Yongin-si 17104, Gyeonggi-do, Republic of Korea

² Department of Industrial & Management System Engineering, College of Engineering, Kyung Hee University, 1732 Deokyoung-daero, Giheung-gu, Yongin-si 17104, Gyeonggi-do, Republic of Korea

³ Integrated Education Institute for Frontier Science & Technology (BK21 Four), Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si 17104, Gyeonggi-do, Republic of Korea

* Correspondence: jayum@khu.ac.kr (J.U.); yunjheo@khu.ac.kr (Y.J.H.)

Abstract: Various accidents caused by alcohol consumption have recently increased in prevalence and have become a huge social problem. There have been efforts to identify drunk individuals using mobile devices; however, it is difficult to apply this method to a large number of people. A promising approach that does not involve wearing any sensors or subject cooperation is a markerless, vision-based method that only requires a camera to classify a drunk gait. Herein, we first propose a markerless, vision-based method to determine whether a human is drunk or not based on his or her gait pattern. We employed a convolutional neural network to analyze gait patterns with image augmentation depending on gait energy images. Gait images captured through a camera allow a complex neural network to detect the human body shape accurately. It is necessary for removing the background behind human shape in the gait image because it disrupts the detection algorithm. A process of conversion into gait energy images and augmenting image data is then applied to the dataset of the gait images. A total of 20 participants participated in the experiment. They were required to walk along a line both with and without wearing the Drunk Busters Goggles, which were intended to collect sober and drunk gait images. Validation accuracy for the recognition of a drunk state in 20 persons was approximately 74.94% under optimal conditions. If the present approach fulfills its promise, we can prevent safety accidents due to alcohol, thus decreasing its burden on industries and society.

Keywords: gait energy image; convolutional neural network; image augmentation; drunk identification



Citation: Park, S.; Bae, B.; Kang, K.; Kim, H.; Nam, M.S.; Um, J.; Heo, Y.J. A Deep-Learning Approach for Identifying a Drunk Person Using Gait Recognition. *Appl. Sci.* **2023**, *13*, 1390. <https://doi.org/10.3390/app13031390>

Academic Editor: Arkady Voloshin

Received: 27 November 2022

Revised: 17 January 2023

Accepted: 17 January 2023

Published: 20 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the World Health Organization (WHO), Korea's alcohol consumption in 2015 was found to be the third largest in East Asia, after Thailand and India [1]. As a result, accidents have occurred in various places, such as drunk driving accidents or accidents caused by alcohol poisoning at industrial sites. In addition, according to the U.S. Department of Labor, 65% of workplace accidents are related to alcohol or drugs [2]. Due to the seriousness of these problems, not only Korea, but the whole world, is paying attention to this situation. Therefore, if it is possible to identify drunk individuals in advance, many alcohol-related accidents may be prevented, thereby minimizing the burden on individuals, industries, and even society due to accidents.

The traditional method of identifying alcohol consumption is the measurement of blood alcohol level using a respiratory system or by drawing blood. However, all of these are direct contact methods, so there is a risk of infection, and they require the subject's cooperation. This is especially true in the case of drawing blood, as it is an invasive method that involves pain. To overcome this problem, a non-invasive method of measuring the

alcohol level in sweat, which is highly correlated with blood alcohol level, through a sensor attached to the skin has recently been proposed [3,4]. However, the method using sweat is available in the only conditions of hot and humid air, exercises, or iontophoresis that accompanies skin irritations. On the other hand, gait analysis is more suitable as a method of identification of a drinker because it can be measured in a no-contact, non-invasive manner without the subject's effort or cooperation.

Traditionally, commercial devices such as motion capture systems, Force Plate, and GAITRite[®] have been used for gait measurement and analysis [5–7]. This system has high reliability and accuracy, but the experimental environment is limited and expensive, and there is the disadvantage that it requires much time for analysis and installation.

As an alternative to avoid these disadvantages, two main approaches have been proposed: a sensor-based method using wearable sensors and a vision-based method using images [8]. First, the sensor-based method uses data extracted from sensors in a smartphone [9–13], a smartwatch [14], and a shoe [15]. Since a smartphone has various motion sensors and high portability, there have been efforts to detect drunk gait abnormalities. When accelerometer and gyroscope data in a smartphone were used, higher accuracy was achieved compared to accuracy when accelerometer data was used alone [9–12]. Yet, the smartphone's location can cause noise and decrease accuracy [16]. A smartwatch can be a better option due to its various sensors, high portability, and lack of variation in wearing location. When accelerometer, heart rate, and temperature data were collected and analyzed, the accuracy was 98.8% [14]. Although such a multi-modal sensing approach using portable devices showed high accuracy in detecting drunk gait, the requirement of the subject's cooperation in wearing the device is a remaining challenge. In addition, a change in gait due to discomfort caused by wearing devices and errors due to differences in wearing positions or contact may occur. Therefore, this will be difficult to apply to a case such as a factory, in which many subjects need to be monitored.

The markerless vision-based method, on the other hand, is a method of analyzing a gait after taking a frame from an image captured by a video camera [8]. This method has the advantage of being able to acquire gait information in a no-contact manner from a distance without the cooperation of the recognition target [17–25]. Thus, even if a body sensor or wearable device is not attached to the target person, we can measure gait in real-time anytime, anywhere, wherever the camera is installed, making it suitable for judging alcohol consumption regardless of location. However, since image data responds sensitively to the recording environment, it is necessary to select an appropriate data processing method to obtain accurate data in various situations.

In this study, we aimed to identify drunk individuals in places of work in order to decrease workplace accidents. We employed the vision-based method because it can identify a drunk person simply passing in front of a camera, without wearing any sensors. To improve accuracy, we used the GEI method, which is simple, able to perform quick calculations, and is not sensitive to recording conditions. We analyzed GEI using deep learning to identify drunk individuals. The data processing method can be divided into a model-based method that analyzes joint movement and an appearance-based method that utilizes silhouette contours obtained through a background subtraction method. The model-based method has the advantage of low sensitivity to camera direction and illuminance, as well as changes in the user's clothes, but has the disadvantage that the calculation is complicated and takes a significant amount of time. Since rapid data processing is important for applying the present technology to a large number of subjects, a simple and fast appearance-based method of calculation is used. The general appearance-based method has a disadvantage in that it is sensitive to recording conditions such as illuminance and the user's attire, but it can be solved by using the gait energy image (GEI) method, which removes the background and averages each pixel of the binarized images [23,24].

As methods for GEI-processed gait data identification, deep learning models that can automatically identify gait by an algorithm and can learn patterns in data and make decisions without the need for manual feature engineering will be used [26]. Among

many deep learning models, a convolutional neural network (CNN), one type of neural network, is commonly used in the image identification field [27,28]. One of the reasons why CNNs are widely used is that they do not require manual tuning (manual-tuning-free), while legacy object detection algorithms require a series of image processing steps, such as contrast optimization or contouring detection. Thus, CNN is a deep learning model that exhibits the characteristics of easy model training because it has fewer connections and parameters than other traditional feedforward neural networks of a similar network size. In addition, since CNN has been used for gait analysis using non-model-based data and its possibility has been verified, this study intends to analyze gait by using it in this study [23].

However, CNN requires a significant amount of data, because it is difficult for it to run normally when the training dataset is not large. To solve this problem, IA (image augmentation) will be used [29]. In Sale's study attempting to implement an algorithm to identify people in images, an experiment was conducted by increasing the quantity of data using the IA technique make utilization of the CNN model feasible [30]. The IA method used in this study included rotation (0–180°), tilting, magnification, brightness, and left/right inversion, and through this, the dataset of 200 people was increased by about 100 times. Afterward, as a result of CNN's deep learning analysis, the accuracy improved to about 96.23% after using IA compared to 82% when IA was not used, which was similar to the accuracy when using a dataset of 800 without using IA. Therefore, it can be confirmed that a performance similar to that of the analysis when using a large dataset is possible using IA, even if the quantity of data is relatively large. Thus, this study also intends to determine the effect of improving accuracy by using IA.

In order to analyze not only the identification using deep learning, but also the unique characteristics of a drunk gait, we will extract parameters and compare the differences in gait characteristics according to whether or not a person has consumed alcohol. Here, three spatio-temporal parameters, such as stride length, stride time, and stride velocity, will be selected as parameters that can represent characteristics well, and these will be compared and analyzed. Our contributions to this paper are the following:

- (1) Detection of drunkenness with image-based gait pattern recognition is attempted;
- (2) Time, length, and velocity of stride are proposed as the features to detect drunken gaits;
- (3) The accuracy of the proposed algorithm shows that the average and standard deviation are 73.94% and 2.81, respectively.

Section 2 describes the materials and methods used in this study. Section 3 demonstrates the results of the accuracy of identification of drunk individuals, the analysis of gait parameters, and their correlation to the accuracy. In Section 4, we discuss the results and findings. We reach a conclusion in Section 5.

2. Materials and Methods

Drunken gait pattern analysis is measured in three main steps: (1) First, when a pedestrian passes by the space where the camera is installed, the background of the camera image is removed in order to accurately input pedestrian information. (2) However, since there are not enough dataset images from which to learn, even with this removed background, similar images are generated for possible situations from various angles by rotating or inverting images. These additional generated images are used for training of the convolution neural network's architecture. (3) The results analysis step is intended to evaluate whether the learned results can recognize drunken walking at either a normal or a valid level. The first result analysis is intended to measure the percentage of the gait that is correctly matched to each pedestrian's situation, and the accuracy of this estimate. Secondly, it performs an analysis to determine highly correlated factors of drunken patterns. The entire procedure is shown in Figure 1.

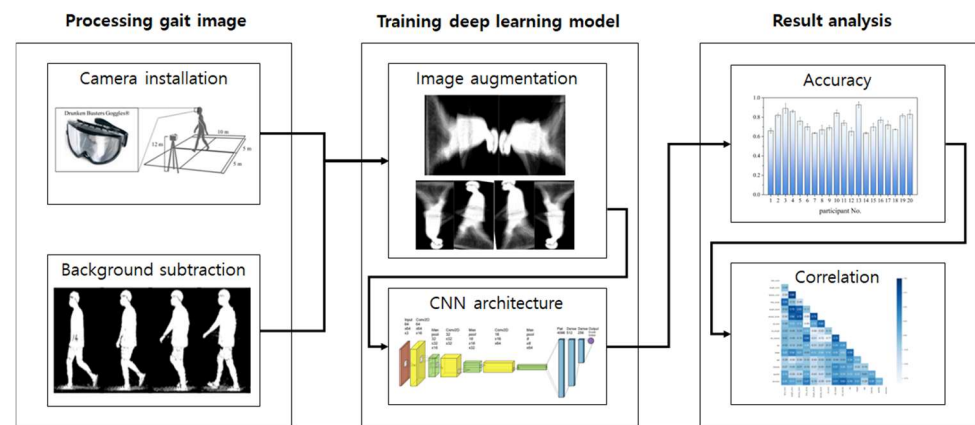


Figure 1. The overall procedure of gait analysis for detecting a drunk pattern.

2.1. Experimental Design

2.1.1. Subjects

To ensure the safety of the experiment, the subjects of this study had no history of cardiovascular or nervous system diseases, and participation was limited to adults with drinking experience. In other words, only those who were able to drink under the Juvenile Protection Act (born in 2002 or earlier, as of 1 January 2021) were subjected to the study.

Twenty subjects were recruited, and the average age of the subjects was 26.1 years. All subjects agreed in writing after being informed of the purpose, method, expected risks, compensation, and right of waiver before they participated in the experiment. This study was conducted after obtaining prior approval from the Kyung Hee University Bioethics Review Board (IRB) before the start of the study (IRB approval number: KHGIRB-21-287).

2.1.2. Setup

The experiment was divided into a questionnaire survey phase and a walking phase. Personal information, such as gender, age, and drinking habits, was collected through a survey.

The subjects' gaits were recorded 10 times each, which consisted of five drunk and five non-drunk walkings. At this time, since there are many limitations to measuring walking in an actual drunken state, goggles (Drunken Busters Goggles®, Drunk Busters of America LLC, Brownsville, TX, USA) that simulate the drinking state were used. Drunk walking consists of walking while wearing the goggles, and non-drunk walking is defined as normal walking without wearing goggles.

Regarding recording, the distance that the subjects walked was 10 m, and the distance between the camera and the floor was 1.2 m. The distance between the camera and the subject was 5 m, and the camera was placed in the center of the 10 m walking distance. The actual walking distance recorded by the camera was about 4 m, which comprised the middle section. To extract only natural human walking, the first and last 3 m were not included in the recorded range. To minimize noise caused by shadows, a shaded area with an illuminance of 3000 lux and which was free from obstructions was selected as the recording location. When recording indoors, it is often difficult to remove the background due to shadows on the wall and sudden changes in lighting; thus, we filmed outdoors in the shade. Regarding the camera, the rear camera (UHD 30 fps) of a mobile phone (Galaxy S21+) was used, and the resolution was 3840×2160 .

2.2. Gait Analysis

This section describes an image processing process for utilizing image data collected from the experiment settings described above for gait analysis. It proceeds through the stages of background subtraction, gait energy image, and image augmentation.

2.2.1. Background Subtraction

During image-based object detection, the recognition algorithm sometimes confuses the target object with the background because the image processing contains an analysis of the color difference and the contour lines from the raw data. However, ambiguity due to similar colors, shadows, and lighting causes the accuracy of the detection algorithm to be reduced. In order to leave only information regarding walking in the input data, it is necessary to remove the background. This is also called the background subtraction technique. The background subtraction technique consists of four steps. The procedure of background subtraction is as follows. The first step is to convert the raw data to an input format. The second step is to calculate the statistical property whether each frame belongs to the background or foreground. And the result of the detection algorithm is compared with the movement in the next frame. This was required in order to verify the results, namely, what was classified as foreground or background.

Typical background subtraction techniques include GMG (Kalman filters and Gale-Shapley matching) [31], MOG (mixture of Gaussian), and MOG2. First of all, GMG is an algorithm that combines the statistical background image removal method and pixel-by-pixel Bayesian segmentation. In general, GMG has been evaluated to be superior in background removal performance, but there is no significant difference in the background detection degree between MOG2 and GMG. However, there is also the advantage that MOG2 runs faster and is not affected by scene changes due to changes in lighting conditions. Therefore, in this study, MOG2 was selected as the image background removal method.

In addition, the frames required in this study were those in which the background was black and the foreground was white. To this end, the noise was removed by adding a morphological calculation to the existing MOG2. The opening technique was used during the morphology calculation, which is a technique to remove noise on the image by applying the dilation calculation immediately after the erosion is performed.

2.2.2. Gait Energy Image (GEI)

A gait energy image (GEI) is a normal cumulative image acquired over time for the entire gait cycle. This is a representative gray-level image with gait information as normalized pixels of the silhouette. To avoid noise from individual images and to ensure smooth synchronization, GEIs are generated based on a continuous sequence in the gait cycle. In particular, GEI shows several major characteristics of gait, such as operating frequency, temporal and spatial changes of the human body, and general shape of the overall body. To generate an accurate GEI, one or more cycles of continuous walking must be included as shown in Figure 2. In addition, compared to other expression methods, GEIs have the advantage of being less sensitive to noise and more likely to obtain excellent results.



Figure 2. Manufacturing and augmentation process of a gait energy image (GEI). (a) Images were captured from background-subtracted video and combined to add up to a consequential GEI. (b) Gait Energy Image (GEI).

Pixels shown in white have the highest intensity and correspond to body parts (head, torso, etc.) that show little movement during the gait cycle. This mainly corresponds to the static domain of the GEI, which contains information related to body type and posture.

On the other hand, pixels with medium intensity appear gray, and it can be seen that the lower part of the leg and the arm are constantly moving. This region shows how movement occurs during walking, and it is the dynamic region of the GEL.

2.2.3. Image Augmentation (IA)

When using existing deep learning models, the number of subjects varies from tens to thousands, and it can be seen that the higher the number of subjects, the higher the accuracy [30]. Using data augmentation techniques such as image augmentation (IA), reliable results can be derived [32]. This is because IA can improve performance by increasing the size of the deep learning training dataset without acquiring new images [30]. In this study, the gait data of 20 subjects were measured, so IA was used to increase the dataset size and improve reliability, as shown in Figure 3.

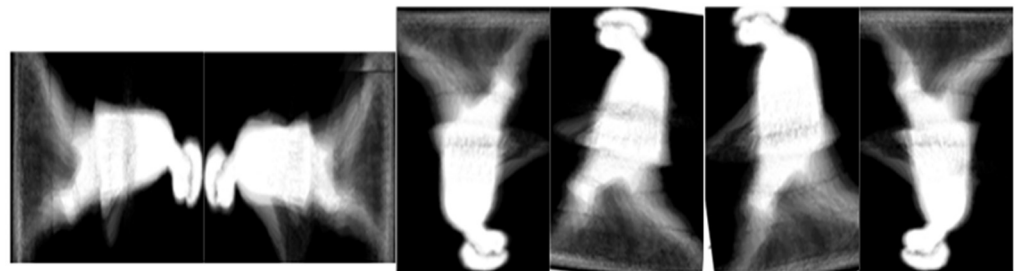


Figure 3. Examples of Image Augmentation (IA) used to enlarge datasets.

3. Results

3.1. Gait Parameter

Stride time, stride length, and stride velocity were selected as parameters for drunk gait, which is distinct from sober gait. These are the parameters that are mainly used in studies related to vision-based abnormal gait evaluation [9]. Walking time and stride length are, respectively, the time and distance taken for one gait cycle. One cycle of gait is from the point when one foot touches the ground to the point when the same motion appears. A total of 10 sober walks and 10 drunk walks were recorded for each subject, and the average value of the measured stride time and stride length for each walk was used as a parameter. The stride velocity was calculated by dividing the stride length by the walking time. As a result of analyzing the parameters extracted from the walking test videos of 20 subjects, we found that there was a significant difference between drunk and sober gait in terms of stride time, stride length, and velocity.

In the drunk gait, the stride time was longer, the stride length was shorter, and the stride velocity was slower than in the sober gait. In addition, as can be seen in Figure 4, this difference was found in all 20 subjects. In the case of walking time, there was a difference of 1 to 8 frames based on a rate of 30 fps.

Furthermore, to comprehensively compare the differences between sober and drunk gaits by parameters, we checked the (drunk/sober) values in Table 1. The drunk gait had a 13% longer stride time, a 13% shorter stride length, and a 22% slower stride velocity than the sober gait.

These differences occurred because the Drunken Busters Goggles caused distortion of vision and decreased gait stability. Since such a decrease in gait stability appears after actual drinking, it is possible to identify a drunk individual by observing a gait that is slower and shorter than usual.

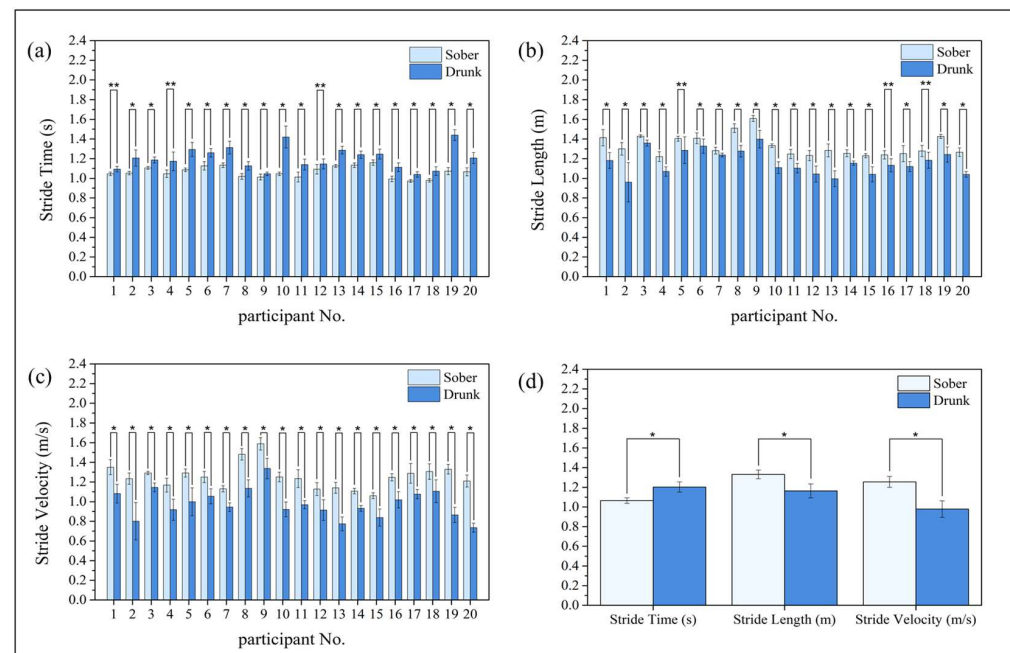


Figure 4. (a) The difference in stride time between sober gait and drunk gait. (b) The difference in stride length between sober gait and drunk gait. (c) The difference in stride velocity between sober gait and drunk gait. (d) The average difference between parameters of sober gait and drunk gait. The difference was statistically significant. (* $p < 0.05$) (** $p < 0.01$).

Table 1. Results of drunk/sober.

Drunk/Sober	Stride Time	Stride Length	Stride Velocity
Average	1.13 ± 0.09	0.873 ± 0.06	0.780 ± 0.08

3.2. Identification

We trained a CNN model using GEI image data to identify drinkers. The diagram of the CNN architecture used in this study is shown in Figure 5. The data consist of 100 images per subject, divided into 10 images for testing and 90 images for training. For each, the drunk data and sober data were set to have the same number. To overcome the disadvantage of a large training dataset being required for a CNN to learn, the data were multiplied through IA.

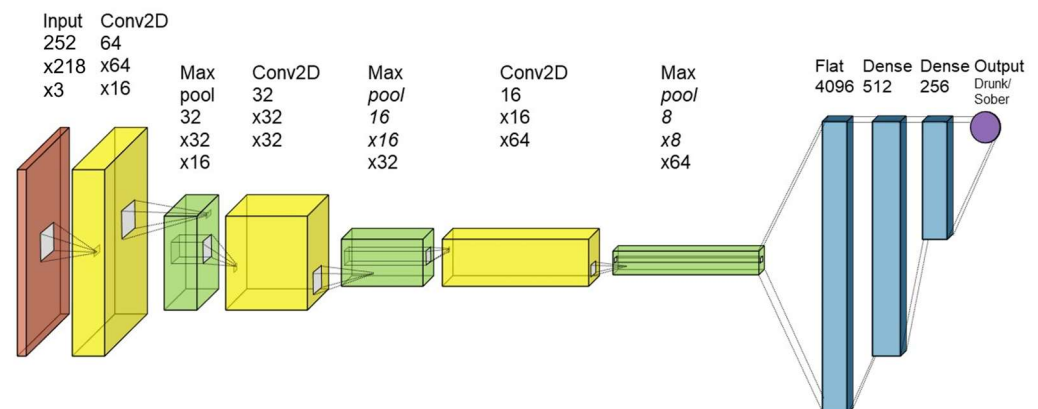


Figure 5. CNN architecture diagram to recognize the drunk gait pattern from GEI image data.

We used scaling, rotation, and flip in IA. In the case of scaling, the range of image enlargement or reduction was set to 0.5 to 1.0 times. The maximum enlargement ratio was

set to 1.0 times so as not to exceed the original, because if a part of the body shown in the GEI image were cut out, accurate identification would become difficult. For rotation, 0–10° rotation to the right and left, as well as 90° and 270° rotation options, were used. For flip, two options were used: up and down and left and right.

In this way, the existing 100 images were multiplied 10 times, so the test data increased to 100 images and the training data to 900 images. Therefore, a total of 1000 images were used for identification through the CNN algorithm.

We optimized the parameters of the CNN model to improve the accuracy of identification of drunk individuals. First, when running CNN, it takes a significant amount of time to train in validation and perform the test because of the large number of datasets. Therefore, epochs were fixed to 200 in order to reduce the computational cost and the time required to run CNN. Then, we increased the dropout rate from 0.3 to 0.9 in steps of +0.2 and checked the accuracy change.

As a result, we found that the accuracy was highest when the dropout rate = 0.3, and decreased as the dropout rate increased; such results corresponded to the well-known characteristic of artificial neural networks. We also checked the identification accuracy according to the kernel size change, but the sensitivity to the kernel size change was negligible.

Therefore, according to these results, the default values of the CNN algorithm parameters were set as epochs = 200, dropout rate = 0.3, and kernel size = 3. The CNN algorithm designed in this way was run with the data from each of the 20 subjects. The resulting loss and accuracy indicating the identification performance are shown in Figure 6.

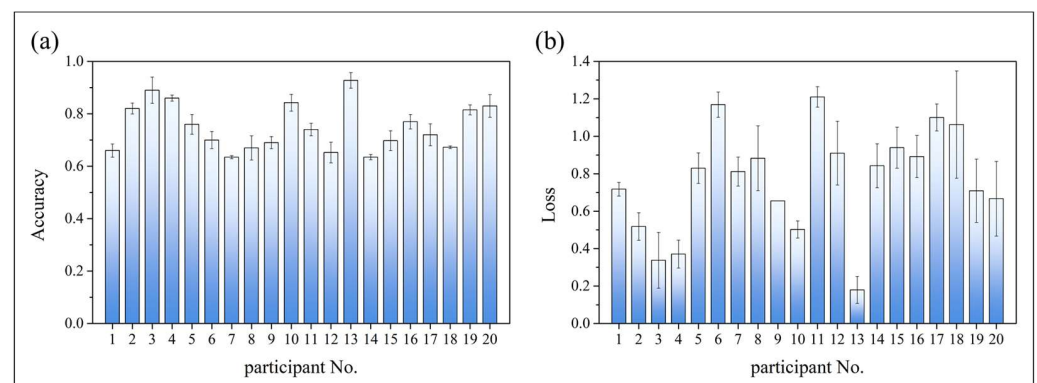


Figure 6. (a) CNN accuracy difference between 20 participants. (b) CNN loss difference between 20 participants.

The identification accuracy for each subject varied from 63.50 to 92.75%. The model showed relatively high identification performance with an average accuracy of 74.93%, and the standard deviation was only 2.81%. Such accuracy was similar to the accuracy with the method using smartphones [9,12]. In the case of loss, the average value was 0.76 and the standard deviation was 0.11. In addition, we conducted the heatmap analysis to evaluate the correlation between accuracy and parameters. Stride velocity showed the highest correlation with accuracy (Figure 7).

The results of the training and validation processes are the following. First, we performed an analysis on 20 people and found that the average accuracy of all passengers was about 74%. The graph from the training result shows that the loss and accuracy trends gradually change and converge, as shown below. Figure 8 is the result of a subject over the course of 200 epochs. The training accuracy of the first subject reached almost 100%, while the validation accuracy increased up to 71%. The validation results were about 74% accurate across all participants, demonstrating that the accuracy tended to be slightly lower than that of the other methods using on-body sensors attached to the subject, such as IMU or the accelerators discussed in previous studies. However, the installation of on-body sensors requires the agreements of the passengers, meaning this type of sensor is not suitable for the purpose of detecting drunken individuals in public spaces.

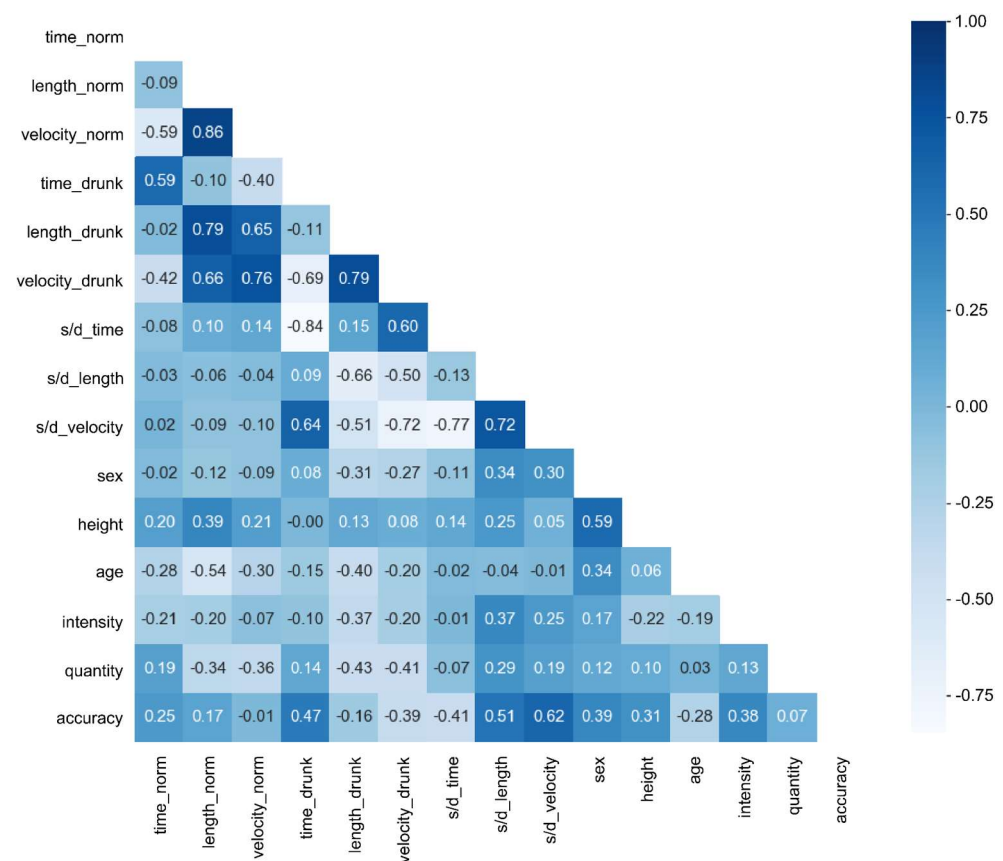


Figure 7. The correlation between accuracy and parameters.

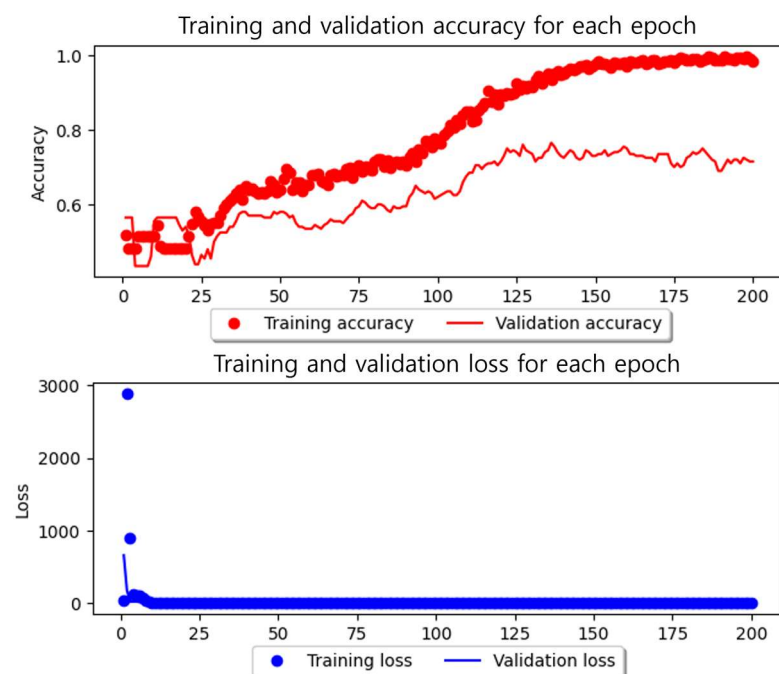


Figure 8. Loss and accuracy of model's training to predict drunken patterns.

4. Discussion

The method of identifying a drunk individual by analyzing a gait video has a great advantage compared to the existing breathalyzer method, in that it can be measured without contact and does not require a device to be attached or worn. However, this

study has a limitation in that the model did not consistently demonstrate high accuracy in the identification of drunk individuals. This problem occurred because the experiment was conducted with a limited number of subjects. Therefore, if a significant amount of training data were to be obtained from a large number of subjects, the accuracy could be improved. In addition, in this study, there is a limitation in that it is possible to identify a drunk individual only by comparing their data with their gait data in a non-drinking state, and thus, identification would be impossible if the non-drinking data were not stored in advance. If the amount of data were to increase, it would be possible to analyze the drunk walking to find the common characteristics of the gait, and then to create a model for identification of drunkenness even without existing data. Another challenging issue is gait recognition based on silhouette segmentation due to the variability in human walking styles and the complexity of the background environment. Silhouette segmentation involves separating the foreground (the person's body) from the background in an image or video. This can be difficult if there are similar colors or textures between the foreground and background, or if the person's pose or clothing is highly variable [33]. Even though we demonstrated promising results with GEI and IA, further training with a larger number of subjects needs to be conducted to illustrate the full potential of the present method.

The methods presented in this study are compared in Table 2. Practically attaching IMU sensors or acceleration sensors to pedestrians in living spaces is not appropriate as a method to monitor drunk pedestrians. Therefore, using an image obtained through a camera is a practical method for this purpose. In addition, regarding the pre-processing method, a remote monitoring system would be required in public places for distributed data processing due to the possible bottleneck of a centralized processing network. This means a small and fast learning model would be better for the remote monitoring system. A lightweight model would be able to determine the state of the individual more quickly. Therefore, the proposed pre-processing method and measurement method is appropriate for detecting drunken persons.

Table 2. Comparison of sensor-based methods and vision-based methods.

	Method	Speed	Raw Data	Attached to A Person	Usage
Sensor-based method	Accelerator on a hand	Fast	Indirect	Need	Personal health management
	Accelerator on a foot	Fast	Direct	Need	Personal health management
	IMU on a foot	Fast	Direct	Need	Personal health management
Vision-based method	Model-based detection	Slow	Direct	Unnecessary	Passenger analysis
	Appearance-based detection	Fast	Direct	Unnecessary	Passenger analysis

In addition, there is a limitation regarding generalization, in that the experiment was conducted outdoors during daylight hours. First of all, this reveals a problem with effectiveness, because alcohol-related problems usually occur after the evening hours. In addition, since the experiment was not conducted in situations where it would be difficult to remove the background, such as in the rain, in shadows, or indoors, it will be difficult to apply it to actual situations in more diverse and complex environments.

Therefore, in a future study, experiments will be conducted at night when the illuminance is lower, indoors, or in various climatic environments, and a method of analyzing gait in combination with other methods, such as gesture detection or body sensors, will be considered. In this study, only CNN, which is a machine learning model commonly used for image analysis, was used, but other machine learning models could also be used for comparisons of accuracy with the intention to build a more suitable machine learning model. Another possible improvement of CNN could be achieved by creating a combined model with time-series learning using recurrent neural networks.

5. Conclusions

Recently, various problems caused by drinking, such as drunk driving accidents and employees working on industrial sites while drunk, are occurring, and the adverse effects of these problems on society are also gradually increasing. There is a need for a technology that can identify drunk individuals and prevent them, as necessary, from sitting in the driver's seat of a vehicle or entering an industrial site. It is difficult to implement this technique with only the traditional breathalyzer method. Therefore, we designed an algorithm to identify drunk individuals by analyzing walking videos and optimized the CNN model to improve its accuracy.

In this study, GEIs were used to identify drunk gait by means of gait image data. The quality of the GEIs varied according to the type of background subtraction method and various experimental conditions. Therefore, we used the MOG2 algorithm, which showed the best background subtraction performance, and set specific experimental conditions to obtain the optimal GEI through repeated experiments. In order to minimize the uncertainty caused by factors such as noise that were expected to occur due to the characteristics of the image data, the illuminance condition was set to 3000 lux, a high-resolution camera was used, and the distance between the camera and the subject was kept constant.

It was also found that the identification performance varied according to the structure of the CNN algorithm and the quantity of data. In particular, because CNN requires a large dataset, a problem occurred because this experiment was conducted with 20 subjects. Thus, the data were augmented using the IA technique. In addition, the accuracy was observed by changing variables such as the dropout rate, kernel size, and epoch. As a result, the best identification performance was achieved with an average accuracy of 74.94% under the conditions of epoch = 32, dropout rate = 0.3, and IA ($\times 10$).

This study is meaningful in that it determined the conditions of the algorithm with optimal performance for identifying drunk individuals. In addition, through analysis between parameters such as gait length, time, and speed, it was discovered that there are various variables in gait data that can identify drunk individuals.

Although this study was limited to identifying drunk individuals using gait image analysis, this technology is expected to be used in other situations where abnormal gait appears in addition to drinking. It can also be applied in various fields, such as the diagnosis of diseases in which abnormal gait is the main symptom, including Parkinson's disease, cancer, and idiopathic normal pressure hydrocephalus. It may also be applied to the detection of abnormal gait due to headaches and dizziness caused by prolonged exposure to harmful substances in environments such as factories or laboratories.

Author Contributions: Conceptualization, Y.J.H. and J.U.; methodology, S.P., B.B., K.K. and H.K.; formal analysis, S.P., M.S.N., B.B., K.K. and H.K.; writing—original draft preparation, S.P., B.B., K.K., H.K., M.S.N. and Y.J.H.; writing—review and editing, J.U. and Y.J.H.; supervision, J.U. and Y.J.H.; project administration, J.U. and Y.J.H.; funding acquisition, J.U. and Y.J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Foundation of Korea, grant number NRF-2021R1A2C1009326. This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant, funded by the Korean government (MSIT), No. RS-2022-00155911, Artificial Intelligence Convergence Innovation Human Resource Development (Kyung Hee University) and Korea Evaluation Institute of Industrial Technology (KEIT) grant funded by the Korea government (MOTIE).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Kyung Hee University (IRB approval number: KHGIRB-21-287; IRB approval date: 16 July 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Alcohol, Total per Capita (15+) Consumption, WHO. Available online: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/total-\(recorded-unrecorded\)-alcohol-per-capita-\(15-\)-consumption](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/total-(recorded-unrecorded)-alcohol-per-capita-(15-)-consumption) (accessed on 18 September 2021).
2. Reily, J. Drug testing & safety: What's the connection? *Occupational Health & Safety*, 1 September 2014.
3. Gamella, M.; Campuzano, S.; Manso, J.; de Rivera, G.G.; López-Colino, F.; Reviejo, A.J.; Pingarrón, J.M. A novel non-invasive electrochemical biosensing device for in situ determination of the alcohol content in blood by monitoring ethanol in sweat. *Anal. Chim. Acta* **2014**, *806*, 1–7. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Kim, J.; Jeerapan, I.; Imani, S.; Cho, T.N.; Bhandodkar, A.; Cinti, S.; Mercier, P.P.; Wang, J. Noninvasive alcohol monitoring using a wearable tattoo-based iontophoretic-biosensing system. *ACS Sens.* **2016**, *1*, 1011–1019. [\[CrossRef\]](#)
5. Bilney, B.; Morris, M.; Webster, K. Concurrent related validity of the GAITRite® walkway system for quantification of the spatial and temporal parameters of gait. *Gait Posture* **2003**, *17*, 68–74. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Webster, K.E.; Wittwer, J.E.; Feller, J.A. Validity of the GAITRite® walkway system for the measurement of averaged and individual step parameters of gait. *Gait Posture* **2005**, *22*, 317–321. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Shin, K.J.; Kang, J.W.; Sung, K.H.; Park, S.H.; Kim, S.E.; Park, K.M.; Ha, S.Y.; Kim, S.E.; Lee, B.I.; Park, J. Quantitative gait and postural analyses in patients with diabetic polyneuropathy. *J. Diabetes Its Complicat.* **2021**, *35*, 107857. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Prakash, C.; Kumar, R.; Mittal, N. Recent developments in human gait research: Parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artif. Intell. Rev.* **2018**, *49*, 1–40. [\[CrossRef\]](#)
9. Arnold, Z.; Larose, D.; Agu, E. Smartphone inference of alcohol consumption levels from gait. In Proceedings of the 2015 International Conference on Healthcare Informatics, Dallas, TX, USA, 21–23 October 2015; pp. 417–426.
10. Kao, H.L.; Ho, B.J.; Lin, A.C.; Chu, H.H. Phone-based gait analysis to detect alcohol usage. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; pp. 661–662.
11. Li, R.; Balakrishnan, G.P.; Nie, J.; Li, Y.; Agu, E.; Grimone, K.; Herman, D.; Abrantes, A.M.; Stein, M.D. Estimation of blood alcohol concentration from smartphone gait data using neural networks. *IEEE Access* **2021**, *9*, 61237–61255. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Zaki TH, M.; Sahrim, M.; Jamaludin, J.; Balakrishnan, S.R.; Asbulah, L.H.; Hussin, F.S. The study of drunken abnormal human gait recognition using accelerometer and gyroscope sensors in mobile application. In Proceedings of the 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, 28–29 February 2020; pp. 151–156.
13. Arora, S.; Venkataraman, V.; Zhan, A.; Donohue, S.; Biglan, K.M.; Dorsey, E.R.; Little, M.A. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Park. Relat. Disord.* **2015**, *21*, 650–653. [\[CrossRef\]](#)
14. Oh, J.K.; Lee, C.Y.; Zhang, B.T. Analyzing multi-sensory data of smartwatch for tipsiness status recognition. In Proceedings of the 2015 Korean Institute of Information Scientists and Engineers (KIISE) Winter Conference, Pyeongchang, Republic of Korea, 15–17 December 2015; pp. 610–612.
15. Park, E.; Lee, S.I.; Nam, H.S.; Garst, J.H.; Huang, A.; Campion, A.; Arnell, M.; Ghalehsariand, N.; Park, S.; Chang, H.J.; et al. Unobtrusive and continuous monitoring of alcohol-impaired gait using smart shoes. *Methods Inf. Med.* **2017**, *56*, 74–82.
16. Silsupadol, P.; Teja, K.; Lugade, V. Reliability and validity of a smartphone-based assessment of gait parameters across walking speed and smartphone locations: Body, bag, belt, hand, and pocket. *Gait Posture* **2017**, *58*, 516–522. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Brunnekreef, J.J.; Van Uden, C.J.; van Moorsel, S.; Kooloos, J.G. Reliability of videotaped observational gait analysis in patients with orthopedic impairments. *BMC Musculoskelet. Disord.* **2005**, *6*, 17. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Liao, R.; Makihara, Y.; Muramatsu, D.; Mitsugami, I.; Yagi, Y.; Yoshiyama, K.; Kazui, H.; Takeda, M. A video-based gait disturbance assessment tool for diagnosing idiopathic normal pressure hydrocephalus. *IEEE Trans. Electr. Electron. Eng.* **2020**, *15*, 433–441. [\[CrossRef\]](#)
19. Lee, L.; Grimson, W.E.L. Gait analysis for recognition and classification. Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 21–21 May 2002.
20. Zhang, Y.; Ogata, N.; Yozu, A.; Haga, N. Two-dimensional video gait analyses in patients with congenital insensitivity to pain. *Dev. Neurorehabilit.* **2013**, *16*, 266–270. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Roh, H. Effect of visual perceptual disturbance on gait and balance. *J. Phys. Ther. Sci.* **2015**, *27*, 3109–3111. [\[CrossRef\]](#)
22. Wang, L.; Tan, T.; Ning, H.; Hu, W. Silhouette analysis-based gait recognition for human identification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1505–1518. [\[CrossRef\]](#)
23. Hu, B.; Guan, Y.; Gao, Y.; Long, Y.; Lane, N.; Ploetz, T. Robust cross-view gait identification with evidence: A discriminant gait gan (diggan) approach on 10000 people. *arXiv* **2018**, arXiv:1811.10493.
24. Wu, Z.; Huang, Y.; Wang, L.; Wang, X.; Tan, T. A comprehensive study on cross-view gait based human identification with deep CNNs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 209–226. [\[CrossRef\]](#)
25. Yoo, H.W.; Kwon, K.Y. Method for classification of age and gender using gait recognition. *Trans. Korean Soc. Mech. Eng. A* **2017**, *41*, 1035–1045. [\[CrossRef\]](#)
26. Khan, M.H.; Farid, M.S.; Grzegorzec, M. Vision-based approaches towards person identification using gait. *Comput. Sci. Rev.* **2021**, *42*, 100432. [\[CrossRef\]](#)
27. Jiang, X.; Wang, Y.; Liu, W.; Li, S.; Liu, J. CapsNET, CNN, FCN: Comparative performance evaluation for image classification. *Int. J. Mach. Learn. Comput.* **2019**, *9*, 840–848. [\[CrossRef\]](#)

28. Sultana, F.; Sufian, A.; Dutta, P. Advancements in image classification using convolutional neural network. In Proceedings of the 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India, 22–23 November 2018; pp. 122–129.
29. Han, D.; Liu, Q.; Fan, W. A new image classification method using CNN transfer learning and web data augmentation. *Expert Syst. Appl.* **2018**, *95*, 43–56. [[CrossRef](#)]
30. Saleh, A.M.; Hamoud, T. Analysis and best parameters selection for person recognition based on gait model using CNN algorithm and image augmentation. *J. Big Data* **2021**, *8*, 1–20. [[CrossRef](#)] [[PubMed](#)]
31. Godbehere, A.B.; Matsukawa, A.; Goldberg, K. Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation. In Proceedings of the 2012 American Control Conference (ACC), Montreal, QC, Canada, 27–29 June 2012; pp. 4305–4312.
32. Veenendaal, A.; Daly, E.; Jones, E.; Gang, Z.; Vartak, S.; Patwardhan, R.S. Drunken abnormal human gait detection using sensors. *Comput. Sci. Emerg. Res. J.* **2013**, *1*.
33. Iwamura, M.; Mori, S.; Nakamura, K.; Tanoue, T.; Utsumi, Y.; Makihara, Y.; Muramatsu, D.; Kise, K.; Yagi, Y. Individuality-preserving silhouette extraction for gait recognition and its speedup. *IEICE Trans. Inf. Syst.* **2021**, *104*, 992–1001. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.