Journal of Marine Science and Technology

Volume 31 | Issue 3 Article 6

Wearable Sensor-Based Walkability Assessment at Ferry Terminal using Machine Learning: A Case Study of Mokpo, Korea

Jungyeon Choi
Division of Navigation Science, Mokpo National Maritime University, Mokpo, Korea, jyc@mmu.ac.kr

Hwayoung Kim
Division of Maritime Transportation, Mokpo National Maritime University, Mokpo, Korea, hwayoung@mmu.ac.kr

Recommended Citation
Choi, Jungyeon and Kim, Hwayoung (2023) "Wearable Sensor-Based Walkability Assessment at Ferry Terminal using Machine Learning: A Case Study of Mokpo, Korea," *Journal of Marine Science and Technology*: Vol. 31: Iss. 3, Article 6.
DOI: 10.51400/2709-6998.2700
Available at: https://jmstt.ntou.edu.tw/journal/vol31/iss3/6

RESEARCH ARTICLE

Wearable Sensor-based Walkability Assessment at Ferry Terminal Using Machine Learning: A Case Study of Mokpo, Korea

Jungyeon Choi [a], Hwayoung Kim [b,*]

[a] Division of Navigation Science, Mokpo National Maritime University, Mokpo, South Korea
[b] Division of Maritime Transportation, Mokpo National Maritime University, Mokpo, South Korea

Abstract

Walkability assessments are becoming more popular, as walking offers numerous health, environmental, and economic benefits to communities. However, previous studies on ferry terminal walkability assessment have been inadequate. This study aimed to develop a wearable sensor system to automatically assess walkability at ferry terminals without conducting surveys. We applied seven machine learning (ML) classifiers to detect different walking environments, including flat ground (FG), downhill slope (DS), uphill slope (US), and uneven surface (UE). The ML models were evaluated across different combinations of classes: 2-class (FG vs. UE), 3-class (U) (FG vs. US vs. UE), 3-class (D) (FG vs. DS vs. UE), and 4-class (FG vs. DS vs. US vs. UE). Among these, support vector machine (SVM) classifiers had the best area under the receiver operating characteristic curves (AUCs) for the 2-class, 3-class (U), and 4-class datasets with 0.970, 0.920, and 0.922, respectively. AdaBoost (AB) performed the best in 3-class (D) with an AUC of 0.953. The least absolute shrinkage and selection operator exhibited better performance in classifying walking environments than maximum relevance and minimum redundancy. This study assessed passenger walkability and improved the built environments at ferry terminals by identifying uncomfortable walking conditions. Furthermore, the results contribute to the development of a passenger walkability evaluation system utilizing intelligent sensors and to the economic revitalization of communities near ferry terminals.

Keywords: Machine learning, Walkability assessment, Gait pattern, Wearable sensor

1. Introduction

Walking is the most common physical activity for preventing adult diseases, maintaining mental and physical health, and enjoying leisure time [1–5]. As walking provides numerous health, environmental, and economic benefits to communities [6–9], a convenient walking environment is required, making planning and designing pedestrian-friendly neighborhoods increasingly imperative. Thus, walkability assessments have become increasingly popular. The walkability of a built environment is measured by its friendliness toward pedestrians and how individuals perceive its quality [6,9].

Numerous studies have evaluated the built environments of transit facilities from the perspective of pedestrians. Shin et al. assessed pedestrian characteristics by measuring the hourly density, walking speed, time, and behavioral patterns of pedestrians at a metro station in Seoul [10]. Kim and Jung demonstrated the feasibility of designing urban parks by considering pedestrian behavioral patterns [11]. Clifton et al. developed the pedestrian environment data scan (PEDS) method to measure pedestrian walking environments [12]. Using PEDS, practitioners could assess pedestrian environments

Received 17 May 2023; revised 20 August 2023; accepted 22 August 2023.
Available online 6 October 2023

* Corresponding author.
E-mail addresses: jyc@mmu.ac.kr (J. Choi), hwayoung@mmu.ac.kr (H. Kim).

https://doi.org/10.51400/2709-6998.2700

for transportation and physical activity, and identify investment priorities. However, there is a lack of research on improving the walking conditions and built environments at ferry terminals. Notably, as Korea is surrounded by sea and has numerous large and small islands, the number of individuals traveling annually to these islands is increasing [13]. Thus, a walkability assessment at ferry terminals is essential.

As mentioned above, only a few previous studies have examined the built environments of ferry terminals and pedestrian convenience. To the best of our knowledge, only one study surveyed passengers about the built environment and pedestrian convenience at a ferry terminal [14]. Although interviews and surveys are commonly used for assessing walkability based on pedestrian perceptions [15], the responses can be biased and lack expert insights. To address this limitation, trained professionals conduct onsite inspections; however, these methods are inefficient [16–18]. Consequently, alternative methods have been proposed to utilize infrastructure-based data such as street view images, videos, or GIS technologies for automated evaluation [10–12], and to use advanced walking pattern recording methods such as capacitive sensor floors or videos with depth sensors [19,20]. Vehicles have been used to identify anomalies in the built environment [21,22]. However, these approaches do not consider pedestrian participation in the evaluation. Owing to recent advances in design and technology, pedestrians can deploy wearable sensors to measure and analyze human physiological responses to their surroundings [22–25]. Furthermore, with the advancements in machine learning (ML), several studies have employed ML techniques. Therefore, it is possible to develop a real-time ML-based walkability monitoring system that integrates individual characteristics by attaching portable and inexpensive wearable sensors to pedestrians.

Therefore, the purpose of this study was to assess the walkability of passengers at ferry terminals. A previous study revealed that passengers felt uncomfortable while walking on a slope and ramp of a ferry during embarkation or disembarkation [14]. Therefore, we conducted an experiment that focused on these walking environments. The primary objective of this study was to develop a system that can automatically evaluate walkability without conducting surveys or direct field assessments. First, we proposed using wearable sensors that use ML classifiers to detect walking environments at ferry terminals. The first objective was to examine the potential use of a wearable sensor to classify walking environments in ferry terminals. We assumed that different walking environments at ferry terminals would change the passengers' gait patterns. In other words, if the walking environment is uncomfortable, it would result in walking characteristics different from those on flat ground. Therefore, if we could classify different walking characteristics on flat ground from those in different walking environments (such as downhill and uphill slopes, and uneven surfaces), we could detect uncomfortable walking areas using wearable sensors. Seven ML algorithms were applied to classify four different walking conditions: flat ground (FG), downhill slope (DS), uphill slope (US), and an uneven surface (UE), such as ferry ramps. The second objective was to identify the best features for assessing walkability. We also hypothesized that variability-related and lateral-directional features are more important for detecting the walking environment at ferry terminals. We applied two feature selection methods: maximum relevance and minimum redundancy (MRMR) and least absolute shrinkage and selection operator (LASSO). We investigated the features selected by each method and found that they performed better in classifying walking environments. The contributions of this study are summarized as follows:

- To the best of our knowledge, the passenger walkability at ferry terminals using wearable sensors is evaluated for the first time.
- The best set of features is made available for future researchers to develop ML models for walking environment classification.
- The evaluation of passenger walkability and improvement of the built environments at ferry terminals by using this basic tool.
- The potential to develop ferry terminals to increase passenger safety and comfort by detecting uncomfortable walking surfaces is enabled.
- The economic revitalization of islands and port areas can be facilitated by improving walkability at ferry terminals, leading to an increase in ferry users.

The remainder of this paper is organized as follows. In section 2, we describe the experimental design and methodology used in this study, including the study area, data collection, data preprocessing, feature selection, ML techniques, and overall implementation. The results are presented in section 3. Section 4 discusses the findings of this study. Finally, in section 5, we present our conclusions and future research directions.

2. Materials and methods

2.1. Experimental site selection

Mokpo city has the largest number of ferry routes between the mainland and the islands. As of 2019, the city accounted for 38.3% of all domestic passengers, recording the largest number of passengers in Korea [13]. Mokpo has 22 ferry companies, 25 ferry routes, and 68 ferry ships carrying passengers, cars, and cargo [14,26]. Therefore, in this study, we selected two sites (north and south ferry terminals) located in Mokpo, Korea, for the experiment, as shown in Fig. 1.

2.2. Participants

Twelve healthy individuals (8 males and 4 females) were included in this study. Participants' information is presented in Table 1. Before the experiment, all the participants read and signed a consent form approved by the Institutional Review Board of Mokpo National Maritime University (protocol code 2022-01-001). The general inclusion criterion was that the participants should be in good health with no difficulty in walking and aged between 19 and 70 years. There were no other exclusion criteria for this study.

2.3. Experimental protocol

A triaxial accelerometer (MetaMotionR, MbientLab, San Francisco, CA, USA) was placed on the waist to collect three-dimensional acceleration data, as shown in Fig. 2. This device was affixed onto the back of each subject's waist using a clipped cover. The sensor location was selected based on previous studies that showed that a waist sensor could reliably predict gait-related test scores and evaluation metrics [27–29]. Moreover, the waist location has the advantage of investigating gait balance abilities [30]. The sampling frequency of the accelerometer was set to 100 Hz. We also used an action video camera (GoPro Hero 8, GoPro Inc., San Mateo, CA, USA) to record the subjects' movements, serving as the gold standard for developing a classification model for different walking segments. They were used to define the environment in which the subjects walked. We captured their walking patterns and environments by observing their movements from behind.

The participants were asked to walk at a comfortable pace at both the north and south ferry terminals. The walking trial in the ferry terminal was divided into three segments: FG, slope, and UE (i.e., the ramp of a ferry ship), as illustrated in Fig. 3. The walking experiment took an average of 2 min

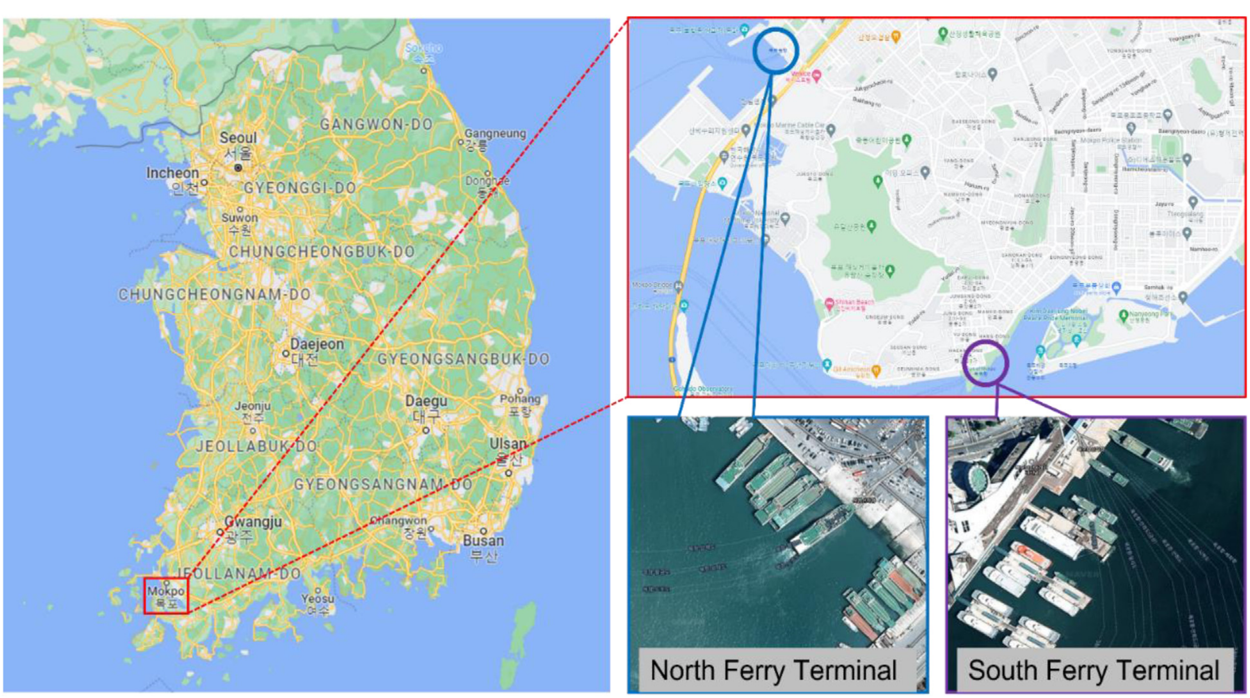

Fig. 1. Experimental sites: the location of north and south ferry terminals in Mokpo, Korea. Note: The map images and pictures were captured from the Google and Naver maps, respectively.

Table 1. Participants' information.

Participant	Gender	Age (years)	Height (cm)	Weight (kg)	Body Mass Index (BMI) (kg/m^2)
1	Male	69	168	65	23.0
2	Female	68	160	65	25.4
3	Female	68	163	63	23.7
4	Female	66	164	60	22.3
5	Male	54	168	75	26.6
6	Male	36	170	68	23.5
7	Male	31	185	72	21.0
8	Male	28	160	60	23.4
9	Female	28	163	55	20.7
10	Male	24	175	67	21.9
11	Male	46	174	72	23.8
12	Male	60	165	60	22.0
Mean (Standard Deviation)		48.2 (18.0)	167.9 (7.3)	65.2 (6.0)	23.1 (1.7)

per person at each terminal, and all participants wore comfortable shoes, such as sneakers.

2.4. Data preprocessing

After collecting the acceleration data from the accelerometer on the waist, we performed several preprocessing steps. First, irrelevant data were removed, and the walking environment was labeled based on the captured videos. To achieve this, we wrote code using MATLAB to eliminate irrelevant data by detecting the starting point within the acceleration data. We manually calculated the walking duration of each class using the recorded video and then automatically labeled the walking environment using the acceleration data. To synchronize the time with the camera system, we manually synchronized the time between the acceleration data and the video by checking the starting point of the experiment. Four classes were labeled as follows: 1) walking on flat ground was labeled as "FG"; 2) walking on a slope was divided into two classes, downhill and uphill, labeled as "DS" and "US", respectively; 3) walking on an uneven surface (a ramp of the ferry) was labeled as "UE". Fig. 4 shows the cleaned and labeled acceleration data.

Subsequently, the detection of step events is necessary to extract gait features. Our previous work described the detection of step events and the extracted features [27,31]. A peak detection method was used to determine the highest peak of vertical acceleration, facilitating step detection. Table 2 lists

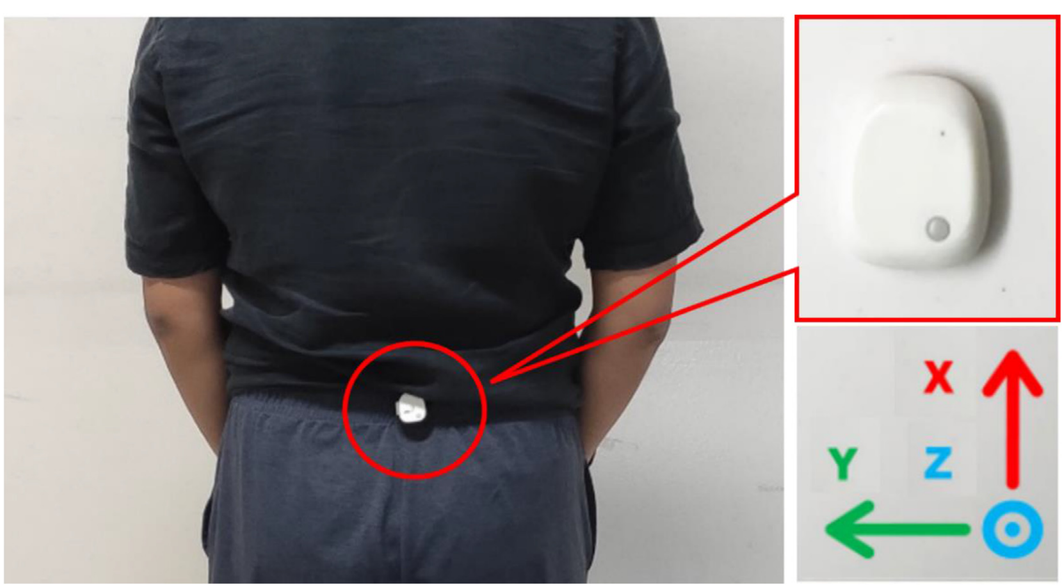

Fig. 2. Sensor location and orientation. An accelerometer was placed on the waist with the X-, Y-, and Z-axis of the sensor representing vertical, lateral, and anterior accelerations, respectively.

Fig. 3. Segmentation of walking trials: flat ground (① and ⑥), downhill slope (②), uphill slope (⑤), and uneven surface (③ and ④). Each subject walked in the order of the numbers.

the twenty gait features. Moreover, each feature was calculated as an average (a), symmetry (s), and variability (v). A total of 60 features were used in this study. The detailed step-detection and feature-extraction methods are described in [27,31], respectively. All features were normalized using zero mean and scaling to unit variance. The normalization equation is as follows:

$$x_{norm} = \frac{x - mean}{sd}, \quad (1)$$

where x_{norm} is a normalized feature, x represents each feature, and *mean*, and *sd* denote the mean and standard deviation values of each feature for all subjects, respectively.

Additionally, data were randomly divided into training (70%) and testing (30%) datasets. The training

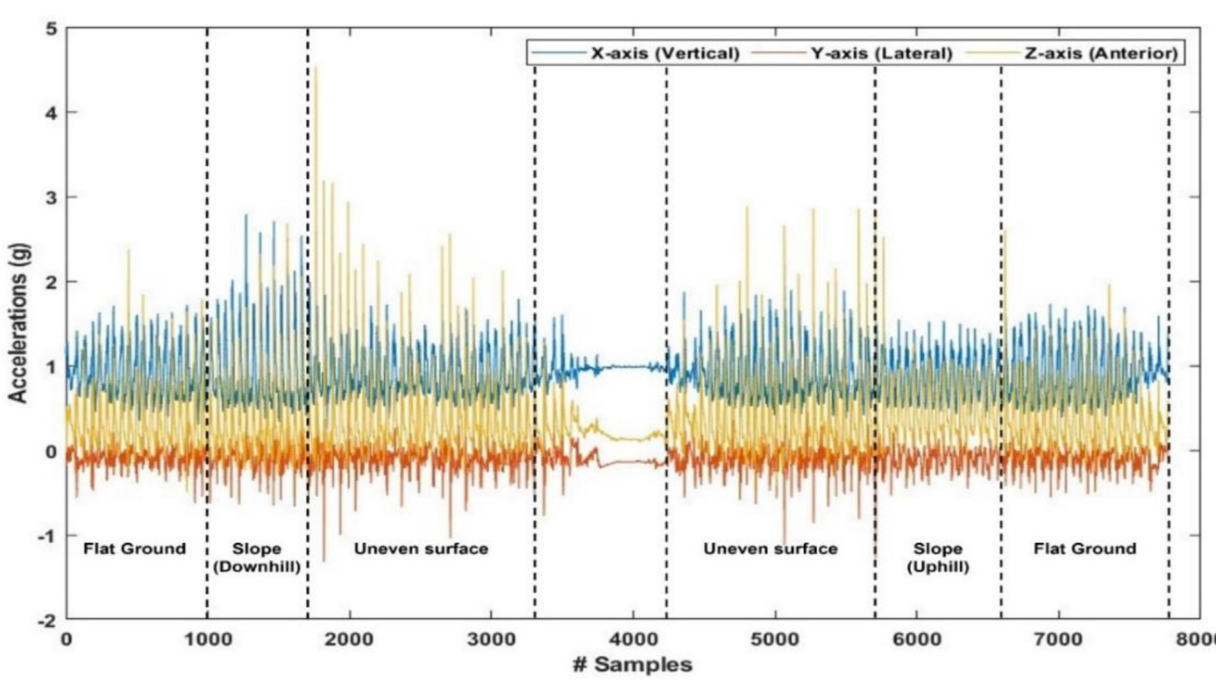

Fig. 4. Example of acceleration data from the waist: each segment was accordingly labeled.

Table 2. List of extracted features [27].

Feature	Description	Mathematical expression
M	Whole step vector magnitude	$\sqrt{L^2+V^2+AP^2}$ for whole step vectors
M10	Initial 10% step vector magnitude	$\sqrt{L^2+V^2+AP^2}$ for initial 10% of step vectors
LM	Lateral vector magnitude during a whole step	$\sqrt{L^2}$ for whole step vectors
VM	Vertical vector magnitude during a whole step	$\sqrt{V^2}$ for whole step vectors
AM	AP vector magnitude during a whole step	$\sqrt{AP^2}$ for whole step vectors
MD	Vector magnitude during double stance	$\sqrt{L^2+V^2+AP^2}$ for ±10% vectors from heel-strike
LMD	Lateral vector magnitude during double stance	$\sqrt{L^2}$ for ±10% vectors from heel-strike
VMD	Vertical vector magnitude during double stance	$\sqrt{V^2}$ for ±10% vectors from heel-strike
AMD	AP vector magnitude during double stance	$\sqrt{AP^2}$ for ±10% vectors from heel-strike
M30	Vector magnitude during mid-stance	$\sqrt{L^2+V^2+AP^2}$ for vectors from 30% of gait cycle
LM30	Lateral vector magnitude during mid-stance	$\sqrt{L^2}$ for vectors from 30% of gait cycle
VM30	Vertical vector magnitude during mid-stance	$\sqrt{V^2}$ for vectors from 30% of gait cycle
AM30	AP vector magnitude during mid-stance	$\sqrt{AP^2}$ for vectors from 30% of gait cycle
LHM	Lateral heel-strike magnitude	max(L) at heel-strike
LHS	Std. of lateral acceleration during initial 10% step	std(L) for initial 10% of step vectors
VHM	Vertical heel-strike magnitude	max(V) at heel-strike
VHS	Std. of vertical acceleration during initial 10% step	std(V) for initial 10% of step vectors
AHM	AP heel-strike magnitude	max (AP) at heel-strike
AHS	Std. of AP acceleration during initial 10% step	std (AP) for initial 10% of step vectors
ST	Step Time	Time between opposite heel strikes

Std.: Standard Deviation, L: lateral acceleration, V: vertical acceleration, AP: anterior-posterior acceleration.

dataset was used for feature selection, ML model fitting, and hyperparameter tuning, while the test dataset was used to evaluate the classification models.

2.5. Feature selection methods

The selection of features for the predictive models is crucial [32]. Irrelevant and redundant features are eliminated to simplify the model. Overfitting can occur when a model includes all possible features. Feature selection aims to prevent overfitting by excluding features that are insensitive to variance sources. Among the various feature selection methods available, we chose MRMR and LASSO, the most common feature selection techniques that have shown effective performance in several existing studies. Furthermore, we conducted several preliminary tests with the initial sample data and found that MRMR and LASSO outperformed other methods. Thus, we examined MRMR and LASSO to determine the most effective feature selection method.

2.5.1. MRMR

Based on the MRMR algorithm proposed by Peng et al. [33], an optimal set of features that are mutually and maximally dissimilar can effectively represent a response variable. This algorithm reduces redundancy and maximizes the relevance of a feature set to the response variable. The algorithm also quantifies the redundancy and relevance of the variables using the pairwise mutual information of features and the mutual information between a feature and its response. The MRMR algorithm calculated the predictor importance scores for each feature, and subsequently, we chose the best features based on a cut-off threshold. Specifically, features were selected if their importance scores exceeded 0.1.

2.5.2. Least absolute shrinkage and selection operator (LASSO)

LASSO was also used to select features relevant to our study. The residual sum of the squares of a vector of regression coefficients was minimized using LASSO [34]. This method reduced the coefficients of less significant variables to zero, resulting in a sparse model. The equation of the loss function in LASSO ($Loss_{lasso}$) is

$$Loss_{lasso}(\beta) = \sum_{i=1}^{n}\left(y_i - \sum_{j} x_{ij}\beta_j\right)^2 + \lambda \sum_{j=1}^{p}\left|\beta_j\right|, \tag{2}$$

where y_i and x_{ij} represent the outcome and variables, respectively, of the i-th subject, n is the number of observations, p is the number of predictors, λ is a non-negative hyperparameter; and β_j is a vector of regression coefficients. When the mean squared error (MSE) is minimum, the best λ is selected using 10-fold cross-validation (CV).

2.6. ML classifiers

In this study, we evaluated seven ML classifiers for categorizing different walking environments in ferry terminals for walkability assessment: decision tree (DT), linear discriminant analysis (LDA), naïve bayes (NB), support vector machine (SVM), k-nearest neighbors (KNN), AdaBoost (AB), and neural network (NN). These models were implemented and tested using a classification learner application in MATLAB R2022a (MathWorks, Natick, Massachusetts, USA). The best hyperparameters for each ML classifier were automatically selected using 50 iterations with the default optimization function of the classification learner app in MATLAB.

2.6.1. Decision tree (DT)

DT uses recursive binary splitting for classification. Each node in a tree is split until the minimum size of the class subgroup or the stop condition is met. The Gini index or entropy is used to assess the quality of each split in a tree [35]. The hyperparameter *MaxNumSplits* was set to 9.

2.6.2. Linear discriminant analysis (LDA)

LDA assumes that the features in each class originate from a multivariate Gaussian distribution with class-specific means and a common covariance matrix. Using these estimates, the discriminant function for each class is calculated. The observations are assigned to the class with the most significant discriminant function values [36]. Notably, this method does not involve any hyperparameters.

2.6.3. Naïve Bayes (NB)

An NB classifier is a simple probabilistic classifier that applies Bayes' theorem while assuming strong independence between features [37]. NB is commonly used owing to its simplicity, tractability, and efficiency [38]. It is suitable for both binary-class and multi-class classification. The hyperparameter *DistributionName* was set to Gaussian.

2.6.4. Support vector machine (SVM)

SVM maps features onto a high-dimensional space using kernels and constructs a hyperplane to effectively separate observations [39–41]. The hyperplane assigns classes based on new observations collected in a high-dimensional space [36]. The hyperparameters *BoxConstraint*, *KernelScale*, and *KernelFunction* were set to 91.0161, 30.6181, and Gaussian, respectively.

2.6.5. K-nearest neighbor (KNN)

The KNN decision boundary was constructed by identifying the k samples that are closest to the observation [42]. In KNN, observations are classified based on a simple majority vote of the k-nearest neighbors [35]. In this study, we used the Minkowski distance metric to calculate distance. The hyperparameter k, which represents the number of neighbors, was set to 14.

2.6.6. AdaBoost (AB)

AB is a statistical classification meta-algorithm that was developed by Freund and Schapire in 1995 [43]. This algorithm can be combined with other learning algorithms to improve performance. The combination of the outputs of different learning algorithms yields a boosted classifier outcome. A weak learner is tweaked to favor instances misclassified by previous classifiers with AB. In some cases, they are less prone to overfitting. AB is commonly used for binary classification but can also be generalized to multiple classes [43,44]. The hyperparameters for AB were set as follows: *NumLearners* = 145, *MaxNumSplits* = 7, and *NumPredictors* = 5.

2.6.7. Neural network (NN)

NN learns by connecting interconnected neurons in a layered structure similar to a human brain. This approach can be used for supervised and unsupervised learning. They are commonly used in speech, vision, and control systems to recognize patterns and classify objects or signals. The NN parameters are determined by weighing it based on its training data and optimized to minimize the prediction error [45]. The hyperparameters for the NN were as follows: *NumLayers* = 1 (i.e., there are two fully connected layers, including the final layer for classification in the NN), *Activation* = sigmoid (the activation function for the final fully connected layer is softmax), *Lambda* = 0.0000002, and *FirstLayerSize* (the neuron size of the first layer) = 116.

2.7. Evaluation metrics

We conducted model training and hyperparameter tuning in the model training phase to minimize the MSE using a 10-fold CV with the training set. MSE was calculated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left[y_{p(i)} - y_{a(i)}\right]^2, \tag{3}$$

where $y_{p(i)}$ and $y_{a(i)}$ are the predicted and actual values, respectively, of the i-th subject for validation, and n is the number of observations.

In the model evaluation phase, the test dataset was used to evaluate the predictive performance of the ML classification models. We evaluated each ML classification model using different class combinations, such as 2-class (FG vs. UE), 3-class(U) (FG vs. US vs. UE), 3-class(D) (FG vs. DS vs. UE), and 4-class (FG vs. DS vs.

US vs. UE) combinations. Several evaluation metrics were used for performance comparison.

Accuracy is the proportion of true positives and true negatives in all cases. The accuracy is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{4}$$

where true positive (*TP*) refers to when a positive class is correctly predicted as positive, true negative (*TN*) refers to when a negative class is correctly predicted as negative, false positive (*FP*) refers to when a negative class is incorrectly predicted as positive, and false negative (*FN*) refers to when a positive class is incorrectly predicted as negative.

In addition to accuracy, we also used the area under the receiver operating characteristic (ROC) curve (AUC) as a model evaluation metric. The ROC curves capture the trade-off between the TP and FP rates [46,47]. This study mainly used AUC for comparing and evaluating different ML classifiers. AUC is a useful measure of classifier performance

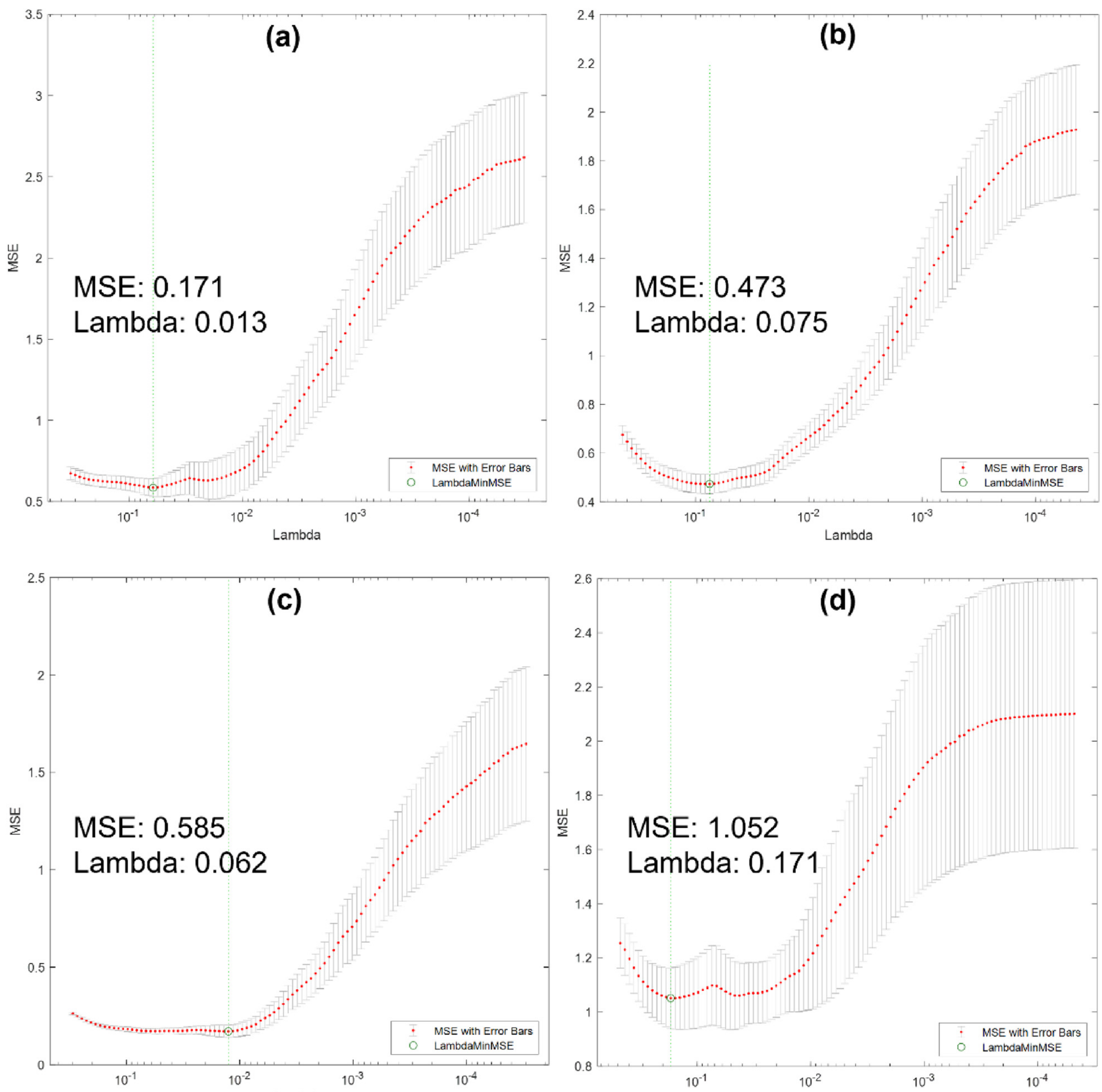

Fig. 5. Parameter λ tuning results for LASSO: (a) 2-class; (b) 3-class(U); (c) 3-class(D); (d) 4-class. The best λ for each model was chosen when the MSE was minimum (λ = 0.013, 0.075, 0.062, and 0.171 for 2-class, 3-class(U), 3-class(D), and 4-class, respectively).

Table 3. Best features selected by MRMR and LASSO for each classification case.

Rank	2-class		3-class(U)		3-class(D)		4-class	
	MRMR	LASSO	MRMR	LASSO	MRMR	LASSO	MRMR	LASSO
1	vLHM	vLHM	vLHM	vLHM	vLHM	vLHM	vLHM	vLHM
2	vVHM	vVHM	vVHS	vVHM	vST	vAHS	vVHS	vVHM
3	sAHS	aST	vM	aST	aLHS	vLMD	vAM30	vLMD
4	–	sAHM	vMD	sAHM	aM30	aST	aST	vAHS
5	–	vAHS	aVM30	sLHM	aST	sLHM	sAHM	sVHM
6	–	sLHM	sLHM	vAHS	–	sAHM	sST	aVHM
7	–	sST	aLMD	sAMD	–	sLMD	aLMD	aST

as it condenses the ROC curve into a value ranging from 0.5 to 1, with 1 representing the best performance [47].

Furthermore, we calculated a confusion matrix to provide detailed classification results. The predicted results are represented using a specific table layout,

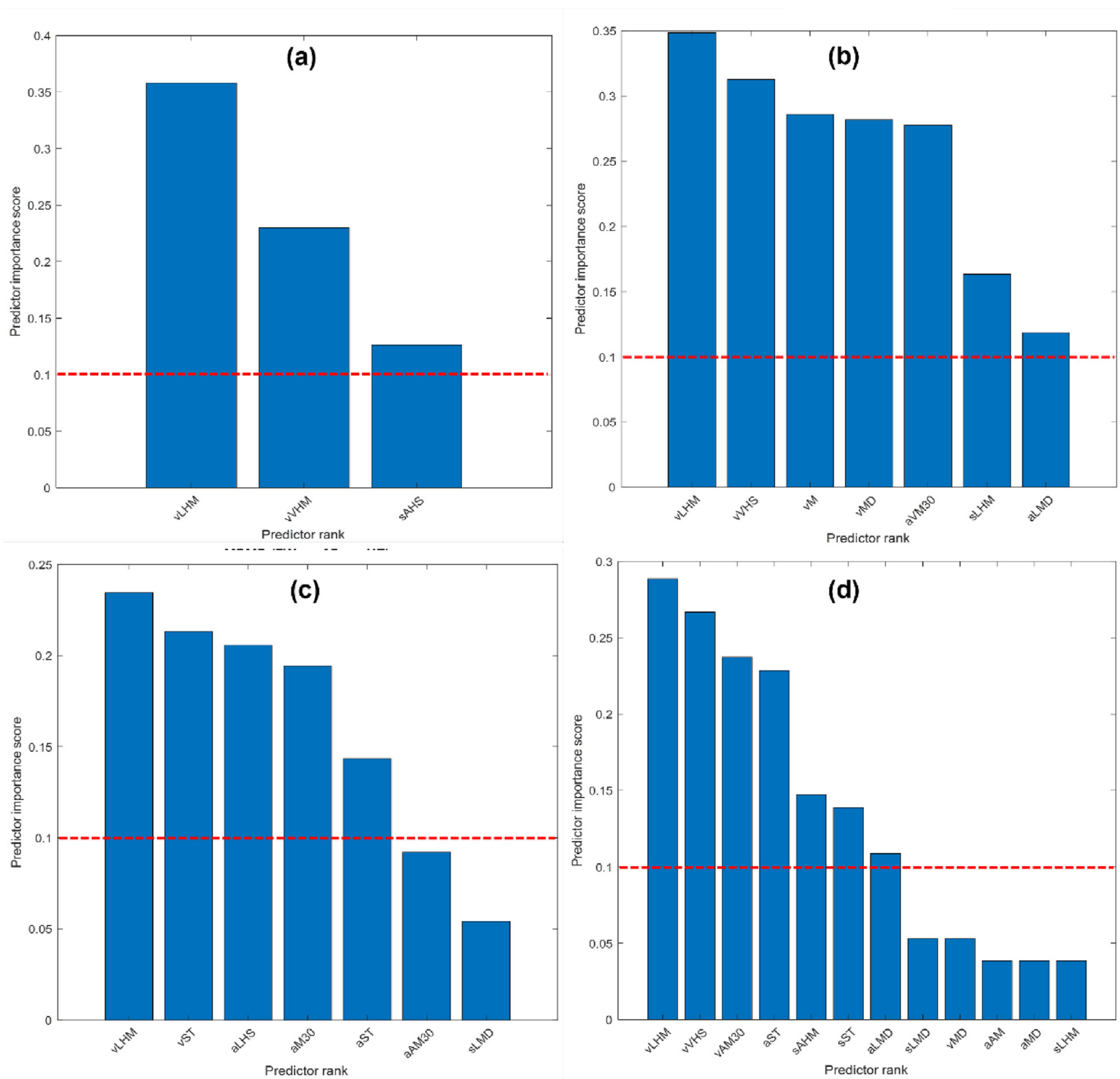

Fig. 6. Selected features by MRMR method for: (a) 2-class; (b) 3-class(U); (c) 3-class(D); (d) 4-class. Red dotted line represents a cut-off threshold (Features were selected when predictor importance score ≥0.1).

facilitating visualization of the ML model performance for binary- or multi-class classification problems.

3. Results

3.1. Results of feature selection

In this study, we employed the MRMR and LASSO methods for feature selection. For the LASSO method, we selected the best tuning parameter λ by finding the smallest MSE through a 10-fold CV. Fig. 5 shows the best parameter λ of the LASSO method for each classification case. The best λ values for different cases were 0.013, 0.075, 0.062, and 0.171, corresponding to minimal MSE values of 0.171, 0.473, 0.585, and 1.052, respectively. Table 3 represents the features selected by LASSO using these λ values. The MRMR algorithm yielded the predictor importance scores for each feature. We chose features with predictor importance scores exceeding 0.1. Fig. 6 and Table 3 present the features selected using MRMR for each classification case. Notably, the feature 'vLHM' was the most important for all cases in both MRMR and LASSO. As expected, variability-related and lateral directional features were selected, as shown in Table 3.

3.2. Classification results

Using the seven ML classification models, we identified the walking environments at ferry terminals using a test dataset. The performance results of the classification for each case are listed in Tables 4–7 with two metrics (i.e., accuracy and AUC). Additionally, we used a confusion matrix to quantitatively analyze the classification accuracy. Fig. 7 displays the confusion matrix of the best classification models based on the AUCs for each case. In the case of 2-class classification, NB, SVM, and NN had the highest accuracy (83.3%), with the SVM performing the best in terms of AUC (0.970). In the case of 3-class(U) classification, both NB and SVM performed the best in terms of accuracy (88.9%) and AUC (0.920). Notably, AB achieved the best performance in terms of both accuracy (83.3%) and AUC (0.953) for the 3-class(D) classification. In the case of 4-class, LDA and NB performed better in accuracy (83.3%), while SVM performed the best in terms of AUC (0.922). Overall, the proposed ML classifiers accurately predicted different walking environments at the ferry terminals in all four classification cases.

Table 4. Classification results for 2-class.

Classifier	Accuracy (%)		AUC	
	MRMR	LASSO	MRMR	LASSO
DT	75.0	75.0	0.79	0.79
LDA	75.0	72.0	0.89	0.86
NB	**83.3**	**83.3**	**0.92**	0.78
SVM	75.0	**83.3**	**0.92**	**0.97**
KNN	75.0	75.0	0.78	0.88
AB	75.0	66.7	0.79	0.85
NN	66.7	**83.3**	0.78	0.83

The values in bold type are the best performance results in accuracies and AUCs for each feature selection method (i.e., MRMR and LASSO) among seven classifiers.

Table 5. Classification results for 3-class(U).

Classifier	Accuracy (%)		AUC	
	MRMR	LASSO	MRMR	LASSO
DT	72.2	77.8	0.796	0.846
LDA	55.6	77.8	0.760	0.883
NB	66.7	**88.9**	0.816	0.913
SVM	**77.8**	83.3	**0.910**	**0.920**
KNN	72.2	66.7	0.843	0.883
AB	72.2	77.8	0.890	0.870
NN	66.7	77.8	0.813	0.890

The values in bold type are the best performance results in accuracies and AUCs for each feature selection method (i.e., MRMR and LASSO) among seven classifiers.

4. Discussion

Several studies have investigated walking environments on sidewalks, roads, city parks, and residential areas [10–12,15,23–25]. However, few studies have examined passenger walking environments at ferry terminals. Only one study focused on the built environment of the ferry terminals [14], but it was survey-based research that could be biased. Thus, this study aimed to assess the walkability of passengers at ferry terminals using a wearable sensor. To the best of our knowledge, this study is the first attempt to assess walking environments at ferry terminals using an accelerometer. We developed ML classification models to detect the different walking environments at ferry terminals in Mokpo,

Table 6. Classification results for 3-class(D).

Classifier	Accuracy (%)		AUC	
	MRMR	LASSO	MRMR	LASSO
DT	72.2	61.1	0.803	0.760
LDA	55.6	77.8	0.683	0.803
NB	72.2	77.8	0.766	0.776
SVM	50.0	72.2	0.743	0.783
KNN	33.3	61.1	0.593	0.710
AB	**77.8**	**83.3**	**0.950**	**0.953**
NN	61.1	61.1	0.756	0.750

The values in bold type are the best performance results in accuracies and AUCs for each feature selection method (i.e., MRMR and LASSO) among seven classifiers.

Table 7. Classification results for 4-class.

Classifier	Accuracy (%)		AUC	
	MRMR	LASSO	MRMR	LASSO
DT	41.7	62.5	0.645	0.787
LDA	54.2	**83.3**	0.815	0.885
NB	58.3	**83.3**	0.805	0.892
SVM	62.5	79.2	0.817	**0.922**
KNN	54.2	70.8	0.872	0.827
AB	**66.7**	70.8	**0.885**	0.902
NN	54.2	54.2	0.775	0.762

The values in bold type are the best performance results in accuracies and AUCs for each feature selection method (i.e., MRMR and LASSO) among seven classifiers.

South Korea. To implement high-performance classification models, we applied two feature selection methods and tuned the hyperparameters of each ML classifier.

Based on the prediction results in Tables 4−7, we found that the accuracies were lower than the AUCs in all cases. The accuracy of a model is measured at specific points and indicates its ability to classify data into appropriate categories based on specific thresholds. However, AUC depends on the ability of the model to demonstrate consistent differences among data with different labels, irrespective of precise data assignment. Accuracy, particularly in this study, is not suitable for performance metrics as even a small change can cause huge differences owing to the small sample size. Thus, we utilized AUC as the main criterion to compare the performance of the models in this study.

The AUC-based results showed that the proposed ML classifiers successfully predicted different walking environments in all cases with high AUCs ranging between 0.920 and 0.970. We found that SVM performed best in 2-class (AUC = 0.970), 3-class(U) (AUC = 0.920), and 4-class (AUC = 0.922), while AB performed best in 3-class(D) (AUC = 0.953). Overall, we recommend using SVM as the most appropriate analysis tool for satisfactory

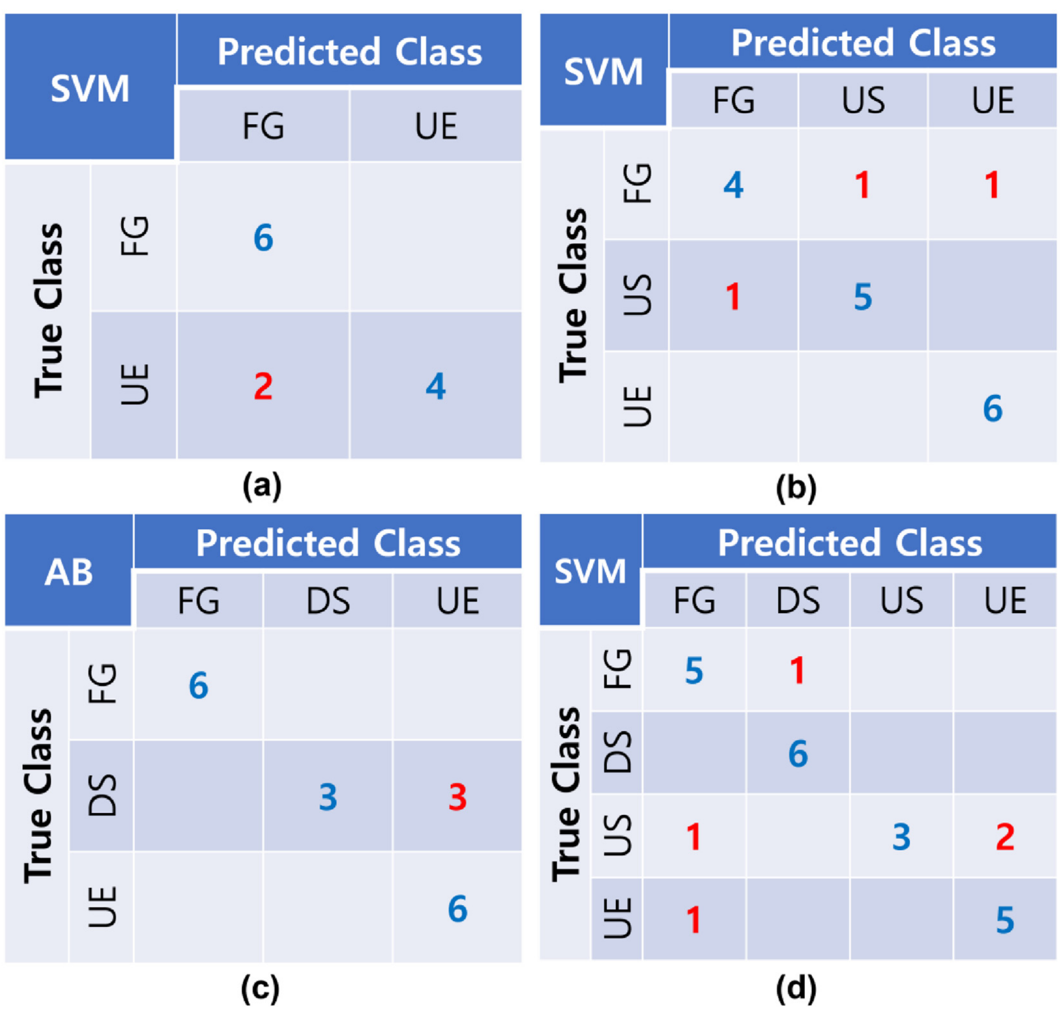

Fig. 7. Confusion matrix of the best classification models for: (a) 2-class; (b) 3-class(U); (c) 3-class(D); (d) 4-class.

passenger walkability perception in ferry terminals. These outcomes support our first hypothesis that different walking environments at ferry terminals affect passenger gait patterns. We detected uncomfortable walking areas using wearable sensors by classifying different walking environments (FG, DS, US, and UE). Thus, it was proven that wearable-sensor-based walkability assessments could be applied at ferry terminals. Based on previous research findings that passengers are uncomfortable walking on slopes and bumpy surfaces such as ferry ramps [14], our proposed models could be used to assess passengers' walkability or improve the built environment at ferry terminals. For example, building a system that automatically monitors uncomfortable sections at a ferry terminal by detecting them using a wearable device can be possible, which can be effectively used to evaluate the walkability at ferry terminals.

Furthermore, this study examined the best set of features for classifying different walking environments. Although the selected features differed slightly depending on the classification case, the 'vLHM' feature was selected as the most important feature in all cases, as shown in Table 3. This result supported our second hypothesis that variability-related and lateral directional features are more important in detecting the walking environment at ferry terminals. This is because, unlike on flat ground, when walking on slopes and uneven surfaces, human bodies can react more from side to side to maintain balance, leading to larger variations in gait patterns. Furthermore, a comparison between MRMR and LASSO showed that the classification models with features selected by LASSO performed better than those using MRMR in all cases (see Table 3). Therefore, we believe that LASSO can be used as a feature selection method for subsequent studies, with the best set of features as primary data to expand the model complexity in future investigations.

Nevertheless, this study had some limitations. First, a limited number of participants were included. Generally, ML classifiers perform better when trained using large amounts of data. Thus, it is difficult to generalize the results obtained from the ML models learned from small samples. However, constraints imposed by COVID in recent years have restricted experiments involving human subjects, resulting in a substantial recruitment effort. Therefore, we need to conduct experiments with more participants in future studies to build a more robust model. Second, the experiment was conducted only at Mokpo. However, this study is the first to evaluate the walkability at a ferry terminal with wearables in the area with the largest number of passengers and ferry routes in Korea. Subsequent studies aim to expand the research target area to include terminals from other regions. Finally, there will be individual differences in walking patterns or characteristics, such as age and sex. However, these individual differences were not considered in this study. Future studies should consider these human factors in the experimental design to understand their potential impact.

5. Conclusions

This study demonstrated that a wearable sensor can be used to detect different walking environments at ferry terminals. The proposed ML classification models were reliable in classifying binary-class (FG vs. UE) and multi-class (3 or 4 classes), including US and DS. We identified uncomfortable walking sections by classifying walking patterns across different walking environments, suggesting that slopes and uneven surfaces should be improved for passengers at ferry terminals. The results showed that SVM performed better for walking environment classification among all ML classifiers. We also examined the best features and a feature selection method for predicting walking environments. The features selected by LASSO performed better in the walking environment classification than those selected by MRMR. We also found that variability and lateral features were associated with walking characteristics on sloped and uneven surfaces.

Our findings can be applied to assess walkability or improve the built environment at ferry terminals by automatically detecting uncomfortable walking environments using a wearable sensor and providing basic data for future studies. The results of this study can also be used to monitor passengers' physical activities by detecting walking patterns at ferry terminals. In particular, building a ferry terminal that prioritizes passenger safety and comfort by identifying uncomfortable walking areas is beneficial for ferry companies and local governments. In addition, enhancing walkability at ferry terminals is expected to increase the number of ferry users, contributing to the economic revitalization of islands and port areas. Nevertheless, further studies with larger sample sizes are required to explore the validity of our findings by applying the proposed methods to different ferry terminals. Furthermore, efforts should be directed toward improving the walkability infrastructure and facilities by using other well-built ferry terminals as benchmarks.

Conflict of interest

None of any author has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Acknowledgments

This research was supported by the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (20220257).

References

[1] Tennøy A, Knapskog M, Wolday F. Walking distances to public transport in smaller and larger Norwegian cities. Transp Res D Transp Environ 2022;103:103169. https://doi.org/10.1016/j.trd.2022.103169.

[2] Rišová K. Questioning gender stereotypes: A case study of adolescents walking activity space in a small Central European city. J Transport Geogr 2021;91:102970. https://doi.org/10.1016/j.jtrangeo.2021.102970.

[3] Rafferty AP, Reeves MJ, McGee HB, Pivarnik JM. Physical activity patterns among walkers and compliance with public health recommendations. Med Sci Sports Exerc 2002;34:1255−61. https://doi.org/10.1097/00005768-200208000-00005.

[4] Pucher J, Dijkstra L. Promoting safe walking and cycling to improve public health: lessons from The Netherlands and Germany. Am J Publ Health 2003;93:1509−16. https://doi.org/10.2105/AJPH.93.9.1509.

[5] Gilmour H. Physically active Canadians. Health Rep 2007;18:45−65.

[6] Ewing R, Handy S. Measuring the unmeasurable: Urban design qualities related to walkability. J Urban Des 2009;14:65−84. https://doi.org/10.1080/13574800802451155.

[7] Frank LD, Sallis JF, Conway TL, et al. Many pathways from land use to health: Associations between neighborhood walkability and active transportation, body mass index, and air quality. J Am Plann Assoc 2006;72:75−87. https://doi.org/10.1080/01944360608976725.

[8] Lee RJ, Sener IN, Jones SN. Understanding the role of equity in active transportation planning in the United States. Transport Rev 2017;37:211−26. https://doi.org/10.1080/01441647.2016.1239660.

[9] Leslie E, Coffee N, Frank L, et al. Walkability of local communities: Using geographic information systems to objectively assess relevant environmental attributes. Health Place 2007;13:111−22. https://doi.org/10.1016/j.healthplace.2005.11.001.

[10] Shin M, Hur J, Kim H, et al. The Characteristics of pedestrians LOS by motion analysis method of the metro transfer passage. The Korean Soc Railway 2016;1:1617−22.

[11] Kim HK, Jung SW. An analysis of the behavior pattern of user for open space in the urban park-focused on the Seoul forest. J Archit Inst Korea 2010;26:37−44.

[12] Clifton KJ, Livi Smith AD, Rodriguez D. The development and testing of an audit for the pedestrian environment. Landsc Urban Plann 2007;80:95−110. https://doi.org/10.1016/j.landurbplan.2006.06.008.

[13] Korea Shipping Association (KSA). Statistics yearbook of coastal shipping. Seoul, Korea: Korea Shipping Association (KSA); 2020.

[14] Kim H, Choi J, Nam Y, Youn J-H. Characteristic Analysis of the Built Environment of Ferry Terminals: A Case Study of Mokpo, South Korea. Sustainability 2022;14:2115. https://doi.org/10.3390/su14042115.

[15] Kim H, Ahn CR, Yang K. A people-centric sensing approach to detecting sidewalk defects. Adv Eng Inf 2016;30:660−71. https://doi.org/10.1016/j.aei.2016.09.001.

[16] Besenyi GM, Kaczynski AT, Wilhelm Stanis SA, Vaughan KB. Demographic variations in observed energy expenditure across park activity areas. Prev Med 2013;56:79−81. https://doi.org/10.1016/j.ypmed.2012.10.011.

[17] Floyd MF, Spengler JO, Maddock JE, et al. Park-based physical activity in diverse communities of two U.S. cities. Am J Prev Med 2008;34:299−305. https://doi.org/10.1016/j.amepre.2008.01.009.

[18] Han B, Cohen D, McKenzie TL. Quantifying the contribution of neighborhood parks to physical activity. Prev Med 2013;57:483−7. https://doi.org/10.1016/j.ypmed.2013.06.021.

[19] Hoffmann R, Brodowski H, Steinhage A, Grzegorzek M. Detecting Walking Challenges in Gait Patterns Using a Capacitive Sensor Floor and Recurrent Neural Networks. Sensors 2021;21:1086. https://doi.org/10.3390/s21041086.

[20] Prochazka A, Schatz M, Tupa O, et al. The MS kinect image and depth sensors use for gait features detection. In: 2014 IEEE international conference on image processing (ICIP). Paris, France: IEEE; 2014. p. 2271−4.

[21] Varona B, Monteserin A, Teyseyre A. A deep learning approach to automatic road surface monitoring and pothole detection. Personal Ubiquitous Comput 2020;24:519−34. https://doi.org/10.1007/s00779-019-01234-z.

[22] Zang K, Shen J, Huang H, et al. Assessing and Mapping of Road Surface Roughness based on GPS and Accelerometer Sensors on Bicycle-Mounted Smartphones. Sensors 2018;18:914. https://doi.org/10.3390/s18030914.

[23] Yoon J, Chun J, Kim H. Investigating the relation between walkability and the changes in pedestrian policy through wearable sensing. Sustainability 2020;12:1−20. https://doi.org/10.3390/su122410447.

[24] Kim H. Wearable Sensor Data-Driven Walkability Assessment for Elderly People. Sustainability 2020;12:4041. https://doi.org/10.3390/su12104041.

[25] Kim J, Ahn CR, Nam Y. The influence of built environment features on crowdsourced physiological responses of pedestrians in neighborhoods. Comput Environ Urban Syst 2019;75:161−9. https://doi.org/10.1016/j.compenvurbsys.2019.02.003.

[26] Pham TQM, Lee G, Kim H. Toward Sustainable Ferry Routes in Korea: Analysis of Operational Efficiency Considering Passenger Mobility Burdens. Sustainability 2020;12:8819. https://doi.org/10.3390/su12218819.

[27] Choi J, Parker SM, Knarr BA, et al. Wearable Sensor-Based Prediction Model of Timed up and Go Test in Older Adults. Sensors 2021;21:6831. https://doi.org/10.3390/S21206831.

[28] Choi J, Knarr BA, Gwon Y, Youn J-H. Prediction of Stability during Walking at Simulated Ship's Rolling Motion Using Accelerometers. Sensors 2022;22:5416. https://doi.org/10.3390/s22145416.

[29] Choi J, Knarr BA, Youn J-H. The Effects of Ship's Roll Motion on the Center of Mass and Margin of Stability During Walking: A Simulation Study. IEEE Access 2022;10:102432−9. https://doi.org/10.1109/ACCESS.2022.3208876.

[30] Van Criekinge T, Saeys W, Hallemans A, et al. Trunk biomechanics during hemiplegic gait after stroke: A systematic review. Gait Posture 2017;54:133−43. https://doi.org/10.1016/j.gaitpost.2017.03.004.

[31] Choi J, Youn J-H, Haas C. Machine learning approach for foot-side classification using a single wearable sensor. In: 40th international conference on information systems, ICIS 2019 (40th international conference on information systems, ICIS 2019). Munich, Germany: Association for Information Systems; 2019.

[32] Tang J, Alelyani S, Liu H. Feature Selection for Classification: A Review. In: Data classification: algorithms and applications. Boca Raton, FL, USA: CRC Press; 2014. p. 37−64.

[33] Peng Hanchuan, Long Fuhui, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern

Anal Mach Intell 2005;27:1226–38. https://doi.org/10.1109/TPAMI.2005.159.
[34] Tibshiranit R. Regression Shrinkage and Selection Via the Lasso. J Roy Stat Soc B 1996;58:267–88. https://doi.org/10.1111/J.2517-6161.1996.TB02080.X.
[35] Gareth J, Daniela W, Trevor H, Robert T. An introduction to statistical learning: with applications in R. Berlin/Heidelberg, Germany: Springer; 2013.
[36] Hastie T, Friedman J, Tibshirani R. The elements of statistical learning. New York, New York, NY: Springer; 2001.
[37] Chen S, Webb GI, Liu L, Ma X. A novel selective naïve Bayes algorithm. Knowl Base Syst 2020;192:105361. https://doi.org/10.1016/j.knosys.2019.105361.
[38] Hand DJ, Yu K. Idiot's Bayes?Not So Stupid After All? Int Stat Rev 2001;69:385–98. https://doi.org/10.1111/j.1751-5823.2001.tb00465.x.
[39] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018.
[40] Ben-Hur A, Horn D, Siegelmann HT, Vapnik V. Support vector clustering. J Mach Learn Res 2001;2:125–37.
[41] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on Computational learning theory - COLT '92. New York, New York, USA: ACM Press; 1992. p. 144–52.
[42] Cunningham P, Delany SJ. k-Nearest Neighbour Classifiers - A Tutorial. ACM Comput Surv 2022;54:1–25. https://doi.org/10.1145/3459665.
[43] Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J Comput Syst Sci 1997;55:119–39. https://doi.org/10.1006/jcss.1997.1504.
[44] Hastie T, Rosset S, Zhu J, Zou H. Multi-class AdaBoost. Stat Interface 2009;2:349–60. https://doi.org/10.4310/SII.2009.v2.n3.a8.
[45] Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 1982;79:2554–8. https://doi.org/10.1073/pnas.79.8.2554.
[46] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res 2002;16:321–57. https://doi.org/10.1613/jair.953.
[47] Fawcett T. An introduction to ROC analysis. Pattern Recogn Lett 2006;27:861–74. https://doi.org/10.1016/j.patrec.2005.10.010.