Research Article

Modeling of Child Stress-State Identification Based on Biometric Information in Mobile Environment

Tae-Yeun Kim,[1] Libor Měsíček,[2] and Sung-Hwan Kim [1]

[1]National Program of Excellence in Software Center, Chosun University, Gwangju, Republic of Korea
[2]J. E. Purkinje University, Ústí nad Labem, Czech Republic

Correspondence should be addressed to Sung-Hwan Kim; shkimtop@chosun.ac.kr

A technology must be developed to automatically identify extreme stress states of children who cannot properly express their emotions when recognizing dangerous situations, which threaten the safety of children, in real time. This study presents a stress-state identification model for children based on machine learning, biometric data, a smart band for collecting biometric data, and a mobile application for monitoring the stress state of the child classified. In addition, through an experiment comparing a dataset using only voice data and a dataset using both voice and heart rate data, we aimed to verify the effectiveness of the combination of the two biosignal datasets. As a result of the experiment, the SVM model showed the highest performance with an accuracy of 88.53% for the dataset using both voice data and heart rate data. The results of this study presented strong implications for the possibility of automating the stress-state identification of a child, and it is expected that the developed method can be used to take preventive measures for dangerous situations to children.

1. Introduction

Owing to the increasing frequency of incidents and accidents affecting children, social attention to the safety of children has increased [1–3]. Accordingly, there has been an increasing demand for technologies that can determine the state of a child in real time. In addition, because children have limited communication skills when they are under stress, they are not able to discern the type of situation they are facing and are therefore severely affected by stress. Protectors are also unaware of how children feel, and this leads to children becoming unknowingly stressed. It is therefore imperative that children receive ongoing care and protection from their protectors. Such risky situations are revealed only after the event, thereby increasing the levels of anxiety among their protectors.

The instant recognition of dangerous situations that affect children is required in order for their protectors to respond to such situations quickly. Closed-circuit televisions (CCTVs) have been widely utilized to detect situations that are dangerous for children; however, it is practically difficult for protectors of children to constantly monitor CCTV videos to identify dangerous situations instantly. Moreover, CCTVs have limitations because these devices transmit and receive image signals based on several people instead of individuals, and violent incidents and accidents do occur in the blind spots of CCTVs. For this reason, a solution for identifying the extreme stress state of a child in real time should be developed to detect dangerous situations more rapidly.

In addition, a new technology named emotional artificial intelligence has emerged with recent advent of the era of the fourth industrial revolution. In particular, with convergence of cognitive science and information and communications technologies, the rapid development of artificial intelligence-based emotional computing technology has allowed the analysis and interpretation of human emotions. In this regard, the importance of technologies on human computer interaction (HCI) is increasing, and along with development and progress of research on HCI, the research focus has shifted from investigation of computer responses from direct user inputs to computer responses based on the emotional inference or user intention [4].

Biometric data have been utilized in previous studies to determine the stress states of individuals. To evaluate the stress levels of potential targets, Bakker et al. [5], Healey and Picard [6], and Jung and Yoon [7] collected biometric signals from workers, drivers, and elderly people, respectively. Setz et al. [8], Melillo et al. [9], and Kurniawan et al. [10] conducted tests involving the learning capabilities of research participants in experimental environments and identified the stress states of these participants. However, such existing studies focused on only reactive stress-state identification and failed to consider the appropriate factors for children with insufficient linguistic and recognition abilities.

In this regard, this study presents a stress-state identification model for children based on both voice and heart rate (HR) data. Voice data, such as the laughing and crying sounds of a child, are important factors for determining the stress state of a child who has difficulty in expressing his or her emotions [11, 12]. In addition, HR data are biometric data that are most frequently used for stress identification and can be used to recognize a child's state of increased tension [4, 9]. Both sets of data are considered to be appropriate for the stress-state identification of a child because such data can be used to identify the stress state of targets regardless of their linguistic and recognition abilities.

For the stress-state identification of a child, machine-learning models such as Naïve Bayes (NB) classification, decision trees (DTs), and support vector machines (SVMs) are used. These models are known as representative classification algorithms and have been widely applied in various fields, including user prediction based on multimodal data learning and automatic document classification [13–16]. In this study, classification models were used to generate a stress-state classification model based on voice and HR data. The developed model classifies the stress state of a child when it receives new data input.

Furthermore, this study presents a system for obtaining the biometric data of a child, identifying the stress state of the child based on the collected data, and monitoring the identification result. The proposed system includes the following three elements: a smart band, the child's stress-state identification model proposed in this study, and a mobile application (app). First, the smart band is equipped with voice and HR sensors to collect a child's biometric data. The child's stress-state identification model proposed in this study calculates biometric data that are collected to identify the stress state of a child. Finally, the mobile app facilitates the monitoring of the identification result in real time. The proposed system enables protectors of children to recognize and respond to problematic situations for their children in real time.

The remainder of this paper is organized as follows. Section 2 describes the related studies on stress-state identification using biosignal information, and Section 3 presents a model and framework for stress-state identification of a child. Section 4 describes the system application and performance evaluation. Finally, Section 5 presents the conclusions with suggestions for future research directions.

2. Related Research

This section introduces existing stress-state identification and analysis methods that are based on biometric data. Table 1 summarizes these methods.

The stress state of targets was identified in previous studies using various combinations of biometric data, including HRs, galvanic skin response (GSR), electrocardiograms (ECG), electromyography (EMG), and electroencephalography (EEG). Among these, HR data have been used for stress-state identification the most frequently in combination with other types of data, because HR changes can be caused by conditions other than stress [17]. Sun et al. [18] identified the stress state of target subjects who performed physical actions based on the HR, GSR, and accelerometer data. However, accurate GSR data were not obtained from subjects who performed numerous actions because the collection of these data depended more on movements than on other biometric signals.

Voice data, such as the crying sounds of children, are important indicators of emotional expressions [19]. Abou-Abbas et al. [20], Rosales-Pérez et al. [21], and Ruvolo and Movellan [11] proposed models for detecting voice data, including crying sounds of children. In particular, the model developed by Ruvolo and Movellan [11] is similar to the model proposed in this study in that the former can detect crying sounds of children using sounds generated from daily activities in kindergartens.

Moreover, Kurniawan et al. [10] conducted an experiment to identify the stress states of subjects by analyzing words that they said based on voice data. Setz et al. [8] and Melillo et al. [9] conducted a test involving the learning capabilities of subjects in an experimental environment that was established.

As indicated earlier, conditions in previous studies are unlikely to be applied to cases of children who have insufficient linguistic and recognition abilities. This study is different from existing studies using voice data because those studies aimed to predict the presence of a disease from a medical perspective or develop robots from an educational perspective rather than recognize dangerous situations for children.

Hence, in this study, both voice and HR data appropriate for children were applied by considering the limitations of previous studies. In addition, a system for detecting dangerous situations for children is proposed by applying the child's stress-state identification model based on these biometric data.

3. Child Stress-State Identification Model

This section presents the child stress-state identification model developed in this study. Section 3.1 defines the stress status of a child that was mainly discussed in this study. Section 3.2 describes the overall framework of this model. Section 3.3 details the data representation and preprocessing processes. Finally, Section 3.4 introduces the proposed model, which is based on machine learning.

TABLE 1: Summary of related studies.

Objects (topics)	Signals used	Analysis methodologies	References
Automatic identification of stress causes of employees	GSR	Adaptive windowing	Bakker et al. [5]
Detecting real-world driving stress	HR, EMG, respiration	Continuous correlations	Healey and Picard [6]
Multilevel assessment model for monitoring elder's health condition	HR, EEG, ECG	SVM, DT, expectation maximization	Jung and Yoon [7]
Personal health system for detecting stress	GSR	Latent Dirichlet allocation, SVM	Setz et al. [8]
Stress elicitation by examination	HR	Latent Dirichlet allocation	Melillo et al. [9]
Activity-aware mental stress detection (sitting, standing, and walking)	Voice, GSR	DT, SVM, K-means	Kurniawan et al. [10]
	HR, GSR, accelerometer	DT, SVM, Bayes network	Sun et al. [18]
Automatic detection of the expiratory and inspiratory phases in newborn cry signal	Voice	Hidden Markov model	Abou-Abbas et al. [20]
Automatic classification of infant crying for early disease detection	Voice	Genetic selection of a fuzzy model	Rosales-Pérez et al. [21]
Automatic cry detection in early childhood	Voice	Gentle-boost	Ruvolo and Movellan [11]

3.1. Stress-State Definition.

Figure 1 graphically defines the stress state of a child according to the combination of voice data and HR data. Level 1 and Level 3 indicate general situations that children can face, whereas Level 2 and Level 4 indicate other exceptional situations that can occur.

Level 1 refers to a condition defined as stress, in which a child cries and shows a high HR. On the other hand, Level 3 refers to a condition defined as nonstress, in which a child does not cry and shows a low HR. A situation in which a child cries in stress is an example of Level 1, and a situation in which a child is at rest is an example of Level 3.

On the contrary, Level 2 refers to a condition in which the crying sound of a child is recognized despite receiving a low HR state. For example, a target child for research does not cry, while another child next to him or her cries.

Level 4 refers to a condition in which a child shows a high HR but does not cry. For example, a child shows a high HR because of physical activities such as running. Conditions that are associated with Level 2 and Level 4 are also defined as nonstress conditions, as is the case with Level 3.

That is, only Level 1 refers to a condition in which a child is stressed, whereas Levels 2, 3, and 4 refer to conditions in which the child is not stressed.

3.2. Framework.

Figure 2 shows the framework of the child stress-state identification model proposed in this study. First, input data including voice and HR data are used as training data through preprocessing, and the training data are used to generate the stress-state identification model for children.

The generated model is applied to determine the stress state of the child based on new tasting data instead of existing training data. When experimental data are input in the model, the state of the child is determined as stress or nonstress. For experimental data, the input data of voice and HR signals are collected from a smart band, which is a wearable device worn by Child A and used as experimental data through preprocessing.

The identification result is transferred to a mobile app installed on the smartphone of the child's protector in real time. When the stress state of the child is detected from the smart band, the app instantly sends a notification to the protector about the state of the child.

3.3. Data Representation and Preprocessing.

In this study, N instance(s) and $T = \{V_1, V_2, V_3, \ldots, V_N\}$ are used. The elements of T have independent identical distribution. Each instance contains an HR datum and an M voice datum or data, and the ith instance is represented as an input vector ($V_i = \{v_{i0}, v_{i1}, \ldots, v_{iM}\}$, $i = 1, 2, \ldots, N$). The jth set of properties of an input vector in T is represented as $U_j = \{v_{1j}, v_{2j}, \ldots, v_{Nj}\}$, $j = 0, 1, \ldots, M$. It is assumed that U_j and $U_j = \{v_{1j'}, v_{2j'}, \ldots, v_{Nj'}\}$, $j \neq j'$, $j' = 0, 1, \ldots, M$ are independent of each other.

Figure 3 shows the process of generating an input vector graphically. Specifically, the number of beats per minute is used as HR data, which are calculated using an R-R interval between two consecutive heart beats. Here, the peak of the first heart beat is defined as R_1, and the peak of the second heart beat as R_2. When the time interval between R_1 and R_2 is 60 seconds, beats per minute (h_i, $i = 1, 2, \ldots, N$) for the ith input vector in the case of occurrence of heart beats are calculated using (1) [22]. Consequently, h_i is set as the first element (v_{i0}, $i = 1, 2, \ldots, N$) of the ith input vector through normalization according to

$$h_i = \frac{60}{R_2 - R_1}, \tag{1}$$

$$v_{i0} = \frac{h_i - \min(h_i')}{\max(h_i') - \min(h_i')}. \tag{2}$$

Here, $\min(h_{i'})$, $i' = 1, 2, \ldots, N$ refers to the minimum value at h_i, and $\max(h_{i'})$, $i' = 1, 2, \ldots, N$ refers to the maximum value at h_i.
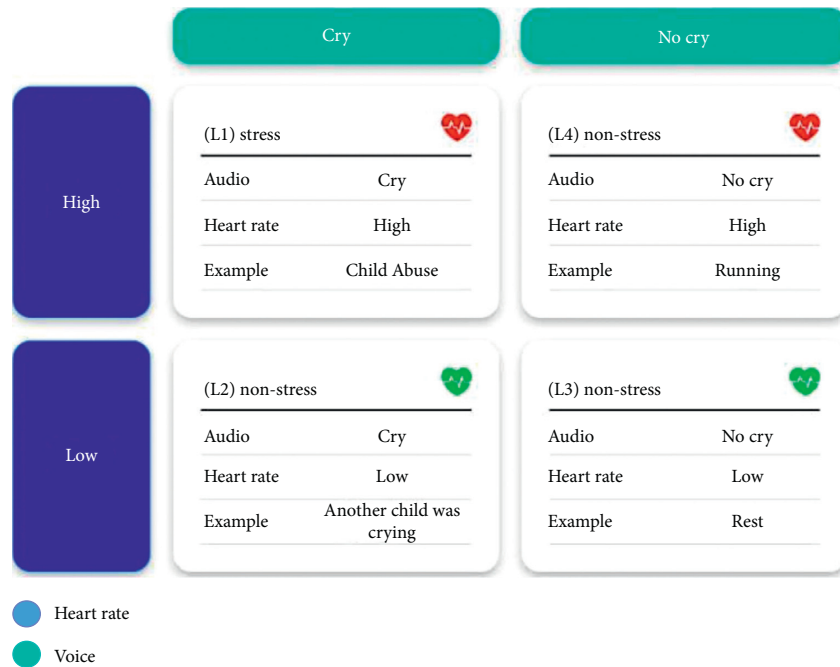
Figure 1: Definition of the stress state of a child according to the combination of voice data and HR data.
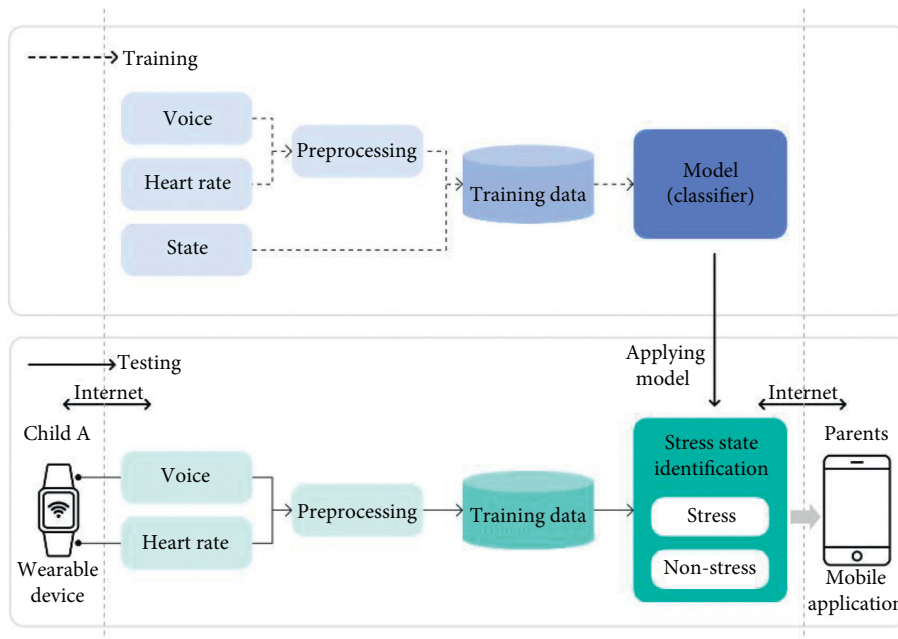

Figure 2: The framework of the child stress-state identification model.

Voice data are feature values of sounds extracted through a series of processes in a sound feature extractor. A sound feature value is obtained through the process of numerically representing sound features through Fourier transformation. When a sound is input, downsampling and preprocessing for windowing are performed. Downsampling is performed to reduce the time required for the extraction of output data, and windowing is performed to prevent sound features from changing discontinuously and rapidly [23].

Finally, sound feature values are normalized to extract ultimate sound feature values.

Sound feature values were extracted using a publicly available sound feature extractor library. Table 2 lists the properties of representative values extracted by the sound feature extractor library used in a study conducted by McKay [24, 25].

In addition, the result of identifying the stress state of a child based on the proposed model and an input vector
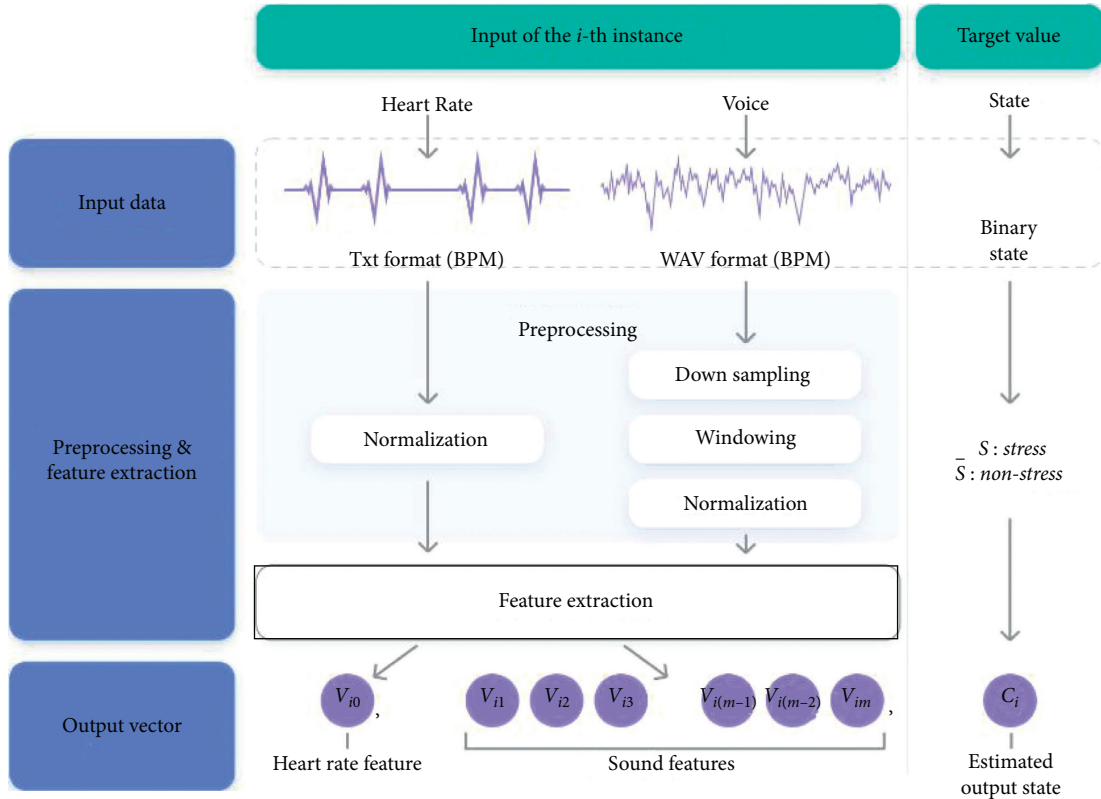
FIGURE 3: The process of generating an input vector graphically.

TABLE 2: Representative sound features.

No	Features
1	Power spectrum
2	Magnitude spectrum
3	Magnitude spectrum peaks
4	Spectral variability
5	Spectral centroid
6	Partial-based spectral centroid
7	Partial-based spectral smoothness
8	Compactness
9	Spectral roll-off point
10	Spectral flux
11	Partial-based spectral flux
12	Method of method
13	Area method of moments
14	Mel frequency cepstral of coefficient (MFCC)
15	Area method of moments of MFCCs
16	Zero crossings
17	Root mean square
18	Relative difference function
19	Fraction of low-energy frames
20	Linear predictive coding
21	Beat histogram
22	Strongest beat
23	Beat sum
24	Strength of strongest beat
25	Strongest frequency via zero crossings
26	Strongest frequency via spectral centroid
27	Strongest frequency via FFT maximum

indicates that C_i, $i = 1, 2, \ldots, N$ is classified into S, indicating stress, and $\overline{S}$, indicating nonstress, as shown in

$$C_i = \begin{cases} S, & \text{if } V_i \text{ is in the "Strees" state,} \\ \overline{S}, & \text{otherwise.} \end{cases} \tag{3}$$

3.4. Stress-State Identification Modeling. The stress state of a child was identified by applying representative classification models based on machine learning, such as NB classification, a DT, and an SVM.

The NB classification model is a type of probabilistic classifiers in which feature values are assumed to be independent, Bayes' theorem is applied for classification into the maximum probability category, and the model has been investigated across a wide range of fields. The NB classification model can be effectively applied to classification of documents or categories, such as spam mail classification [26].

The DT classification model illustrates decision rules using diagrams, classifies the target group into several small groups, and makes prediction based on these small groups [27]. The DT classification model is constructed by establishing split criteria, stopping rules, and evaluation criteria for the purpose of analysis and data structure. Then, the tree undergoes a pruning process to remove improper branches that may increase the classification error. Finally, the success

of the DT is evaluated through cross-validation analysis using the verification data of several diagrams. The DT classification model is used as a method to increase the accuracy of future prediction analysis and to describe the analysis process [28].

The SVM model is a supervised learning model for pattern recognition and is mainly used for classification problems. The distance between the decision boundary and the closest data from the training data is called margin, and the data located at the closest distance to the decision boundary is called the support vector [29]. While using the SVM model, if the two classes to be classified have a separating plane, the classes can be separated by introducing a kernel function. Based on this, Vapnik proposed a classification method using a nonlinear kernel [30]. If the training data provided in a specific space is mapped to a higher dimensional space using a kernel function, a hyperplane with a maximum margin can be obtained based on the characteristics of the data. In other words, with datasets, the decision boundary is nonlinear; however, in a specific high-dimensional space after mapping, linear separation becomes possible.

The NB classification method, which is the first classification model applied in this study, calculates the probability ($P(C_i = S|V_i)$, $i = 1, 2, \ldots, N$) that V_i, $i = 1, 2, \ldots, N$ is in the S condition using (4). It calculates the probability ($P(C_i = \overline{S}|V_i)$, $i = 1, 2, \ldots, N$) that V_i, $i = 1, 2, \ldots, N$ is in the $\overline{S}$ condition based on $1 - P(C_i = S|V_i)$.

$$P(C_i = S|V_i) = \frac{P(C_i) \cdot P(v_{i0}, v_{i1}, \ldots, v_{iM}|C_i)}{P(V_i)} \propto P(C_{i'j}|C_{i'}). \tag{4}$$

HR and voice data, which have continuous values, are utilized through supervised discretization [31]. Moreover, as the independent status of elements of U_j is assumed, $P(v_{i0}, v_{i1}, \ldots, v_{iM}|C_i)$ can be expressed as $\prod_{j=0}^{M} P(v_{ij}|C_{i'})$, $i' = 1, 2, \ldots, N$, $j = 0, 1, \ldots, M$. The NB classification method determines the S condition when $P(C_i = S|V_i) > P(C_i = \overline{S}|V_i)$ is observed in V_i, $i = 1, 2, \ldots, N$ based on the calculated probability. The opposite case is classified as the $\overline{S}$ condition for decision making.

The DT method, which is the second classification model applied in this study, selects properties of an input vector that has the greatest information gain as a decision node and keeps pruning until the ultimate decision value is classified [32]. The information gain is calculated based on entropy, which is used to calculate the complexity of S or $\overline{S}$ in T. The entropy is minimized when only either S or $\overline{S}$ is determined. In other words, it has the minimum value when only one condition is determined, i.e., the condition in which a child is stressed or the condition in which a child is not stressed.

The properties of an input vector that serves as a decision node are expressed as $U_{j^*} = \{v_{1j^*}, v_{2j^*}, \ldots, v_{Nj^*}\}$. To establish a split point in U_{j^*}, which has continuous elements, the mean of two consecutive elements in U_{j^*} is calculated. Subsequently, the entropy of the mean of $N - 1$ is calculated,

and a split point at the lowest entropy is represented as $\text{Split}(v^*)$. In this case, the DT classifies decision values based on $v_{ij} < \text{Split}(v^*)$ and $v_{ij} > \text{Split}(v^*)$ [31].

The SVM method, which is the third classification model applied in this study, classifies the stress status of a child by identifying a decision boundary that maximizes the margin between an input vector belonging to S and an input vector belonging to $\overline{S}$ in a vector space. Equation (5) indicates an objective function.

$$\text{Min} \ \frac{1}{2}w^2 + \lambda \sum_{i=1}^{N} \xi_i,$$

$$\text{s.t.} \ y(w^T \cdot V_i + b) - 1 + \xi_i \geq 0, \quad \xi_i \geq 0, \tag{5}$$

$$\sum_{i=0}^{N} \xi_i \leq \lambda, \quad i = 1, 2, \ldots, N.$$

Here, y refers to the stress state of a child, and it has the value of 1 in S and that of -1 in $\overline{S}$. w denotes a normal vector vertical to a decision boundary. A soft margin method is applied, and the slack variable is placed on the decision boundary to consider the noise of data that cannot be classified based on only S and $\overline{S}$. ξ_i refers to the size of an error allowed in the model, and λ refers to a variable that controls the effects of a decision boundary.

When a discriminant $((\widehat{w}^T \cdot V_i + \widehat{b}) - 1 + \widehat{\xi} = 0))$ is generated through the application of training data, the SVM classifies the stress state of a child by determining whether an input vector belongs to $(\widehat{w}^T \cdot V_i + \widehat{b}) - 1 + \widehat{\xi} > 0$ or $(\widehat{w}^T \cdot V_i + \widehat{b}) - 1 + \widehat{\xi} < 0$.

4. System Implementation and Experiment Results

4.1. System Implementation. In this section, data transfer structure according to terminals is described. Data transfer is classified as data transfer performed between a smart band and a mobile app, and that performed between a mobile app and a server.

A smart band is a customized band for children and is equipped with voice and HR sensors. Biometric data of children are collected through these sensors. With respect to the data collected, the status of the smart band connection and the data collection status can be checked using a mobile app. The collected data is preprocessed to generate an input vector that is transferred to a server and stored in a database (DB). The mobile phone should include a communication module for facilitating data transfer with a smart band agent, and one for facilitating data reception from the server. That is, an agent refers to a smart band measurement device including measurement sensors, and a DB refers to a server in which biometric data are stored. Communication modules including BLE Manager and HTTP Client are installed on the mobile phone.

The agent transfers data to BLE Manager of the communication modules using Bluetooth according to the ISO/IEEE 11073 Personal Health Data (PHD) Standards. HTTP

Client exchanges information with the mobile app and server using HTTP methods for JavaScript Object Notation (JSON) and images (JPG/PNG). The entire resources on the web are represented as URI to connect these resources with each other, and both HTTP methods (e.g., GET, POST, PUT, and DELETE) and URL are used to transfer data [33, 34].

Figure 4 shows a flowchart for biometric data transfer from a smart band to a web server. This chart shows three stages. In the first stage, biometric data measured in a smart band are transferred to a mobile app over a Bluetooth connection. In the second stage, a user checks measurement values on the mobile app and transfers them to a web server via the Internet. In the third stage, the user can search biometric data stored in the DB and monitor them on the screen of a mobile app.

In the first stage in Figure 5, the user selects a smart band that he or she is willing to connect with a mobile app. The smart band selected is registered for Bluetooth pairing. After the Bluetooth pairing process is complete, the smart band is connected to the mobile app and transfers biometric data measured to the app.

In the second stage, biometric data transferred from the smart band are processed to be displayed on the screen of the mobile app for the user. These data are temporarily stored in a local DB of the mobile app and transferred to a web server. Biometric data transferred to the web server are stored in the DB of the server, and the connection is terminated.

In the third stage, biometric data that are stored in the server can be searched, checked, and monitored on the mobile app.

When the app is executed on the mobile phone, a sign-in request screen is displayed. When an automatic sign-in function is checked on the sign-in screen, sign-in information is stored in the app, and the sign-in screen is omitted from that time. When sign-in verification is complete, a smart band list screen is displayed. The user selects a smart band that he or she is willing to use, checks the Bluetooth pairing status for this device, and waits for Bluetooth pairing registration. Subsequently, the Bluetooth service for connection with this device is declared and initialized. When the Bluetooth service is declared, the smart band is searched around the mobile phone and connects when it is in close proximity. Measurement values are received, transferred to a server, and provided in the form of daily and monthly reports.

Figure 6 shows the entire processes of the app. First, the sign-in process is executed, and sign-in information is stored. As shown in Figure 7, an onBind() Callback method should be generated to receive information from the smart band. This method returns an IBinder object, which defines the interface in which clients can interact with services. As shown in Figure 8, it also executes an initialize() function to set the Bluetooth connection.

The initialize() function processes the initial Bluetooth setting and returns a value obtained through such process to a Boolean. Subsequently, the Bluetooth service is loaded from a system service and stored in BluetoothManager. In addition, the basic adapter of the device is accessed by getAdapter() from BluetoothManager and stored in BluetoothAdapter. Using these processes, initial tasks for using the Bluetooth function are completed.

With respect to the null check in BluetoothAdapter, finish() or termination is implemented immediately when the mobile phone of the user does not support the Bluetooth function or when the Bluetooth function is not activated. Basically, when Bluetooth is not supported, getAdapter() returns null and leads to null in BluetoothAdapter. When BluetoothAdapter.isEnabled() is false, this means that the Bluetooth function is turned off. To search surrounding Bluetooth devices, a BluetoothLeScanner object is generated.

As shown in Figure 9, a setFilterDeviceName() function searches surrounding Bluetooth devices based on the name of the Bluetooth device to be used and checks the status of this device registered for Bluetooth pairing. The name of the device to be used is transferred using filterDeviceName, and surrounding Bluetooth devices are searched based on the name transferred through start_Scan(). Using getPairedCheck(), the Bluetooth device that is matched is examined to determine whether it is registered for Bluetooth pairing.

The connect() function requests a connection with a Bluetooth device using the MAC address of the device. The connection result is returned to a Boolean. First, it confirms the declared status of BluetoothAdapter. When it is not declared, it returns false. When declaration is confirmed, information on the Bluetooth device to be connected based on its MAC address is loaded. When this information is not loaded, the function also returns false. When the connection is fully confirmed, it examines details of the GATT server connection with the selected Bluetooth device and data transfer and reception and connects the Bluetooth device.

An onCharacteristicChanged() function receives information from the connected Bluetooth device and processes it to be used. It generates and sends a broadcastUpdate() function to the activity screen of the app. It also stores such information in a remote DB.

Figure 10 shows a screen of the developed app on which a protector can monitor the stress state of the target child over time. This app presents the stress state of a child in real time and according to time zones. It serves as a medium for real-time monitoring, which enables users to apply it habitually and constantly [35, 36].

Specifically, it is operated in a one-to-one connection with the smart band. It displays information that is transferred from the smart band in real time, along with previous information stored in DBs, on a screen. When the stress status of the child is identified, it instantly sends a notification with that status to his or her protector to inform the protector about the situation.

4.2. Experiment Results

4.2.1. Experiment Environment. With respect to the experimental data used to verify the child's stress-state identification model proposed in this study, two types of datasets were established, as indicated in Table 3.

First, in Dataset A, properties of an input vector consist of only sound feature values extracted from voice data.
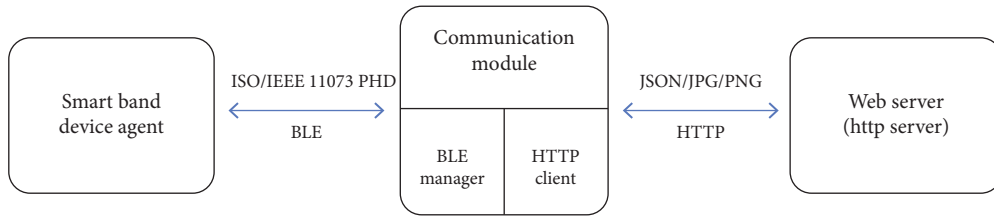
Figure 4: Data transmission structure.
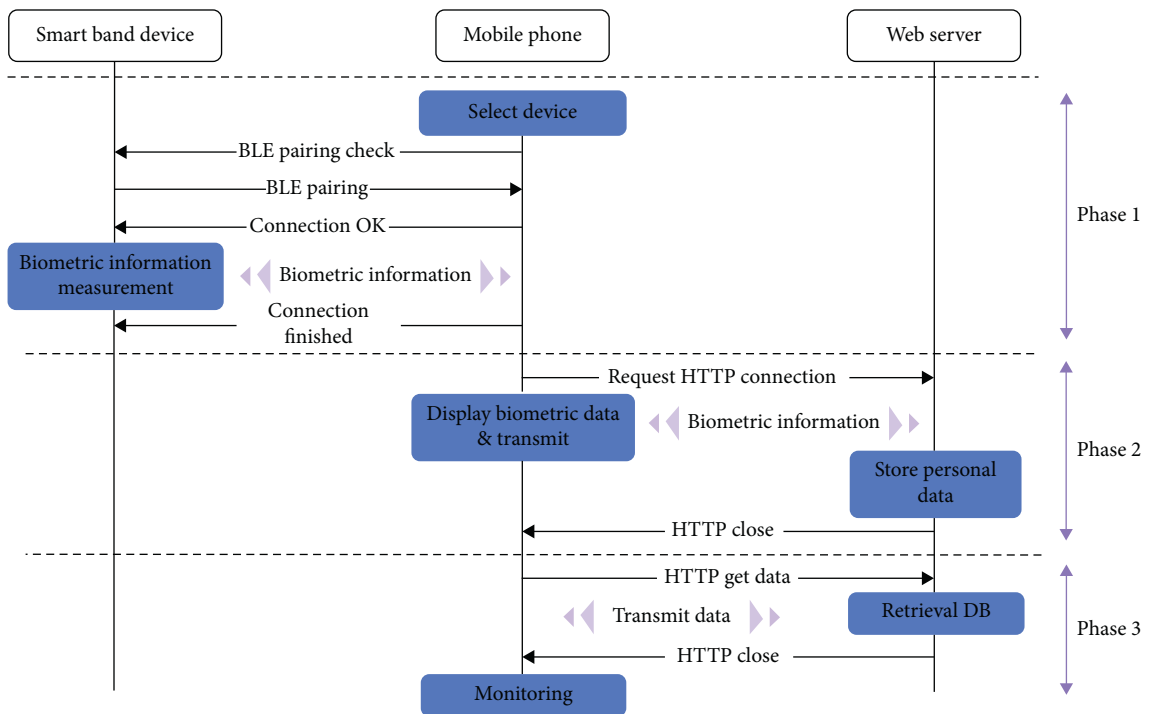

Figure 5: Data flow diagram.
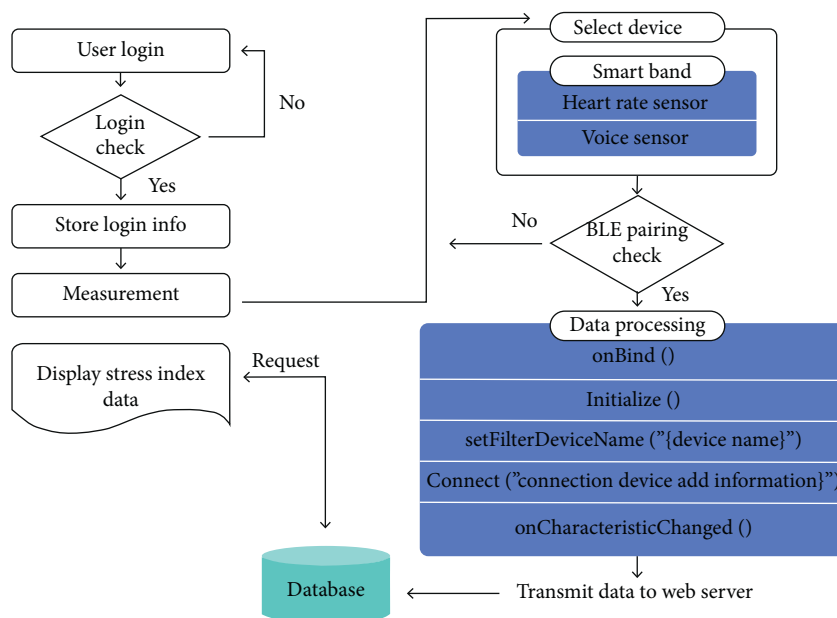

Figure 6: Procedure flowchart.

```
Public IBinder onBind (Intent intent) {
        if (initialize ()) {
                // success
        } else {
                // fail
        }
        return mBinder;
}
```

FIGURE 7: onBind() function.

```
Public Boolean initialize (){
    if (mBluetoothManager == null) {
      mBluetoothManager = (Bluetooth Manager)
getSystemService (Context.BLUETOOTH_SERVICE);
            if (mBluetoothManager == null) {
                    return false;
            }
    }
    mBluetoothAdapter =
        mBluetoothManager.get.Adapter();
    if (mBluetoothAdapter = = null || !mBluetoothAdapter.isEnabled()) {
        broadcatUpdate (BLUETOOTH_OFF_STATE);
        return false;
    } else {
        bondedDevice = mBluetoothAdapter.getBondedDevices();
    }
    mBLEScanner = mBluetoothAdapter.getBluetoothLeScanner();
    If (mBLEScanner = = null)
            return false;
    }
    return true;
}
```

FIGURE 8: initialize() function.

```
Public void setFilterDeviceName (String filterDeviceName) {
        this.filterDeviceName = filterDeviceName;
            // Search for nearby devices
        start_Scan();
            // Calling an existing pairing check method
        getPairedCheck();
}
```

FIGURE 9: setFilterDeviceName() function.

However, as Dataset B uses both voice and HR data, properties of an input vector contain both voice feature values and HR feature values in this dataset.

Publicly available voice data included various sounds and sound effects based on records having a duration of approximately 10 s [37]. Data representing the stress state included crying sounds, and data representing the nonstress state included laughing sounds, conversation sounds, screaming sounds, clapping sounds, and noise from daily activities. With respect to the setting for voice data extraction, a sample extraction rate of 16 kHz was set, and the size of the window was set as 32 ms by referring to existing studies. To minimize data loss, the window overlapping was established as 66% [38, 39].

The HR data of six normal subjects were collected for one week during which they performed daily activities. Accordingly, we conducted a survey based on these subjects and classified the HRs of these subjects in both stress and
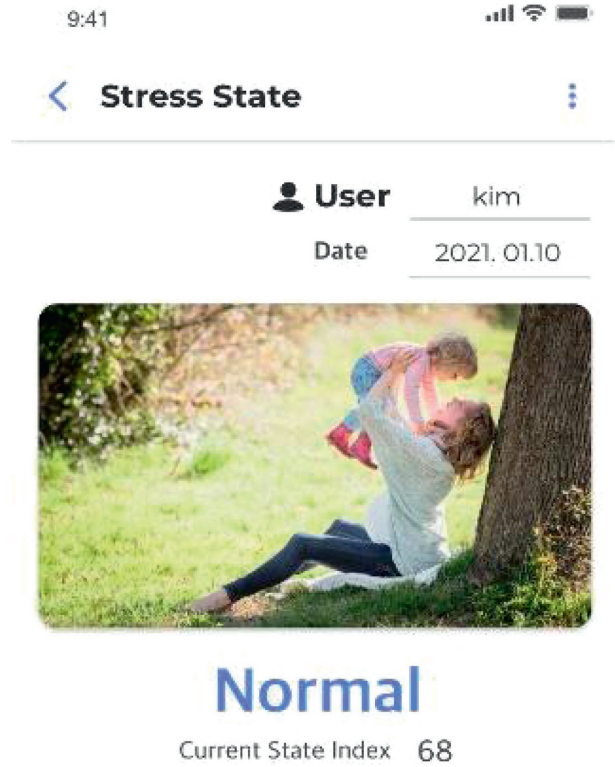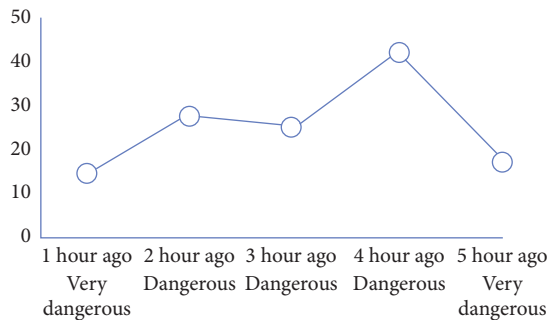

FIGURE 10: The stress-state monitoring application.

TABLE 3: Dataset description.

Datasets	Features
Dataset A	Sound features
Dataset B	Sound features and heart rate feature

nonstress states according to the survey result. Then, HR values that are appropriate for each voice based on voice data collected were compared with HRs of subjects, and were formed according to expert advice. The total number of input vectors in Dataset A and Dataset B is 992; the number of input vectors classified to be in the S condition is 250, and the number of input vectors classified to be in the $\overline{S}$ condition is 742.

Parameter setting for classification models is as follows. With respect to the DT model, pruning was adopted to

prevent the spread of unnecessary branches. The threshold for controlling pruning was established as 0.5. With respect to the SVM model, a linear kernel was used, and the value of λ was set to 1.0.

Both the general accuracy and balanced accuracy were used as indices for evaluating the performance of the proposed model considering data imbalance. Balanced accuracy can detect biased classification in an imbalanced dataset. It becomes algebraically consistent with general accuracy when a dataset stays in balance [40, 41].

In this study, the general accuracy refers to the ratio of the amount of data in which the state of a child is correctly classified according to the total amount of data, as shown in

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}. \quad (6)$$

Here, True Positive (TP) indicates the result in which the case of a stressed child is correctly classified as the S condition, and True Negative (TN) indicates the result in which the case of a child who is not stressed is correctly classified as the $\overline{S}$ condition. However, False Positive (FP) indicates the result in which the case of a child who is not stressed is incorrectly classified as the S condition, and False Negative (FN) indicates the result in which the case of a stressed child is incorrectly classified as the $\overline{S}$ condition.

The balanced accuracy is calculated using (9) based on the sensitivity obtained from (7), which refers to the ratio of data in which the case of a stressed child was practically classified as the S condition, and the specificity obtained through (8), which refers to the ratio of data in which a case of a child not stressed was practically classified as the $\overline{S}$ condition.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (7)$$

$$\text{Specifictiy} = \frac{\text{TN}}{\text{FP} + \text{TN}}, \quad (8)$$

$$\text{Balanced Accuracy} = \frac{(\text{Sensitivity} + \text{Specifictiy})}{2}. \quad (9)$$

The ratio of training data to experimental data is $9:1$. In addition, 10-fold cross-validation was performed to enable the entire dataset to be used in the experiment.

4.2.2. Experiment Results and Analysis. Table 4 lists the number of TP, TN, FP, and FN cases used in the accuracy calculation. It was found that the number of TP cases in which the case of a stressed child was correctly predicted was very low under the condition of Dataset A applied. An analysis revealed that this result was obtained because data state values were established in a biased manner by assuming exceptions in the $\overline{S}$ condition. Such biased data state values led to data imbalance. In particular, the result of the DT model under the condition of Dataset A was derived because of the limitation of this rule-based model. In other words, it developed branches that are based on a variable that had the highest value among variables with similar quantity of information, thereby leading to a limited number of TP cases.

On the contrary, the DT model and the SVM model, respectively, showed 179 and 188 TP cases under the condition of Dataset B. The number of TP cases of these models under the condition of Dataset B increased significantly compared to those under the condition of Dataset A. This result indicates that HR data were located at the upper decision node in the DT model and increased the performance of this model, and that the decision boundary of the SVM model also became more precise because of HR data. That is, it was verified that the increased number of TP cases under the condition of Dataset B affected the model accuracy improvement, and that the combined use of voice and HR data was appropriate for the identification of the stress state of a child.

Table 5 shows the sensitivity, specificity, general accuracy, and balanced accuracy of classification models according to datasets. The experimental result showed that the mean general accuracy and balanced accuracy of the three classification models were calculated to be 64.69% and 52.31%, respectively, under the condition when Dataset A was applied, and 80.58% and 77.98%, respectively, under the condition when Dataset B was applied.

It was confirmed that the classification models exhibited a lower balanced accuracy than general accuracy. This result was derived because the balanced accuracy was calculated based on the sensitivity and specificity and thus facilitated the detection of errors caused by class imbalance. In addition, the DT and SVM models exhibited a low balanced accuracy under the condition when Dataset A was applied. Given that the sensitivity was close to 0, it was determined that these models classified most instances as the $\overline{S}$ condition.

The experimental result verified that the classification models exhibited a better performance when both voice and HR data were used. Based on this result, it can be inferred that a supplementary method is needed to identify the stress state of a child, i.e., one that cannot be determined based on only voice data. It was also determined that the combination of voice and HR data overcame the limitation of the use of only voice data and increased the performance of the developed model.

Then, improvements of the general accuracy and balanced accuracy of the SVM model, which exhibited the most excellent performance under the condition of Dataset B, were confirmed through comparisons with the other models to compare the performance of the three classification models used in this study. The accuracy improvement ratio is calculated using (10). Here, $\text{Accuracy}_{\text{SVM}}$ refers to the general accuracy and balanced accuracy of the SVM model, and $\text{Accuracy}_{\text{comp}}$ refers to the general accuracy and balanced accuracy of the model compared with the SVM model.

$$\text{Improvements} = \frac{\text{Accuracy}_{\text{SVM}} - \text{Accuracy}_{\text{comp}}}{\text{Accuracy}_{\text{comp}}}. \quad (10)$$

Under the condition of Dataset B, the general accuracy of the SVM model increased by 1.39% and 34.38% compared to that of the DT model and the NB model,

TABLE 4: Numbers of correct and incorrect classification instances.

	Dataset A			Dataset B		
	Naïve Bayes	Decision tree	Support vector machine	Naïve Bayes	Decision tree	Support vector machine
True Positive (TP)	180	13	8	182	179	188
True Negative (TN)	430	718	595	446	612	615
False Positive (FP)	312	24	147	296	130	127
False Negative (FN)	70	237	242	68	71	62

TABLE 5: Classifier performance according to dataset.

	Sensitivity		Specificity		Accuracy		Balanced accuracy	
	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
NB	78.82	73.24	58.32	60.03	60.51	65.88	65.87	68.12
DT	64.38	70.84	95.54	93.18	73.83	87.32	48.93	81.88
SVM	63.53	81.56	74.36	93.92	59.73	88.53	42.13	83.95
Average	68.91	75.21	76.07	82.38	64.69	80.58	52.31	77.98

respectively, while the balanced accuracy improved by 2.53% and 23.24%, respectively. As indicated in this result, the general accuracy of the SVM model was improved significantly compared to that of the NB model. On the contrary, that of the SVM model improved only slightly compared to that of the DT model. This result was derived because both DT and SVM models classified the stress states of children based on high probabilities under the condition of Dataset B. However, the DT model exhibited higher performance than the SVM model under the condition of Dataset A. Given this, it is determined that the combination of voice and HR data enabled the SVM model to exhibit a better performance than the DT model in terms of general accuracy.

Table 6 and Figure 11 show the improvement rates of general accuracy and balanced accuracy of the three classification models under the condition of Dataset A compared to those under the condition of Dataset B.

The experimental result indicated that the three classification models exhibited a better performance under the condition when Dataset B was applied than under the condition when Dataset A was applied. It was found that the SVM model showed the highest accuracy improvement. The general accuracy and balanced accuracy of this model increased by 48.22% and 99.26%, respectively. In particular, its balanced accuracy increased significantly. This result implies that the combination of voice and HR data reduced the error caused by the imbalanced dataset observed in Dataset A.

However, the general accuracy and balanced accuracy of the NB model were calculated to be approximately 60% under the conditions of both datasets, thereby being significantly low. This result was obtained because the datasets used in this study contained a lower number of input vectors than that of properties of input vectors. From the experiment, it was found that most probabilities of each property of instances were approximated to be zero, and that the probability of an input vector being included in a condition through continuous multiplication converged to zero. Thus, it is determined that the logarithmic transformation method should be applied to conditional probability equations.

TABLE 6: The improvement rates of general accuracy and balanced accuracy according to datasets.

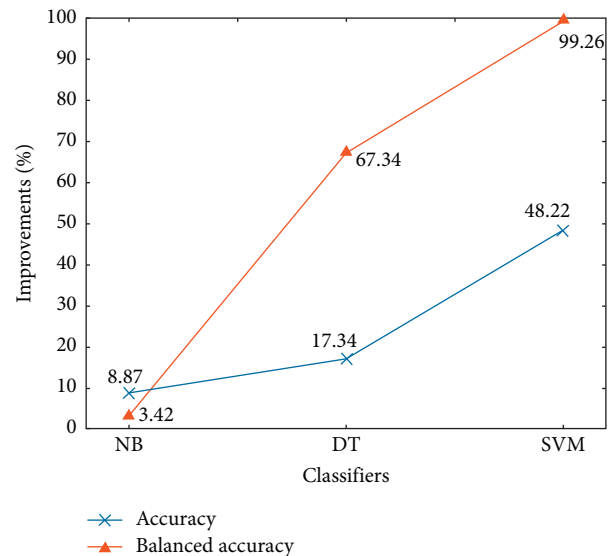|	Accuracy	Balanced accuracy
NB	8.87	3.42
DT	17.34	67.34
SVM	48.22	99.26

FIGURE 11: The improvement rates of general accuracy and balanced accuracy according to datasets.

5. Conclusion

A technology must be developed to automatically identify extreme stress states of children who cannot properly express their emotions when recognizing dangerous situations, which threaten the safety of children, in real time. This study presented a child's stress-state identification model based on biometric data. When a child who cannot easily show his or her expressions encounters a dangerous situation, such a situation tends to be revealed after the child was affected by a

risk, thereby increasing relevant problems. In this regard, this study presented a methodology for identifying the stress state of a child in real time and recognizing dangerous situations for the child. Voice data and HR data, which can represent emotions of children, were used as biometric data, and a child stress-state identification model based on machine learning was developed.

In this study, the stress state of a child is classified into four cases, and we aim to perform the classification using normalized voice data and heart rate data and to verify the reliability of the proposed system. The NB, DT, and SVM models, which are commonly used as classification models for machine-learning-based biosignals analysis, were also used as classification models in this study. In addition, we aim to verify the effectiveness of the combination of the two biosignal datasets through an experiment comparing a dataset using only voice data and a dataset using both voice and heart rate data. As a result of the experiment, in the dataset using both voice data and heart rate data, the results of accuracy for NB, DT, and SVM models were 65.88%, 87.32%, and 88.53%, respectively, with SVM model showing the best performance. From these results, the proposed model showed the highest performance when datasets using both voice and heart rate data were used. Furthermore, this study presented the framework for collecting biometric data using a smart band, identifying the stress state of a child based on the proposed model, and transmitting the state that is identified to a mobile app in real time. These methods enable protectors of children to recognize dangerous situations for children instantly and to respond to these situations promptly.

The results of this study show that the child's stress-state identification model based on two types of biometric data and the framework, both of which were developed in this study, can be practically applied to the detection of dangerous situations for children. It is expected that this model can also be verified in practical conditions as long as an experimental environment based on children can be established through cooperation with their protectors. In addition, it is necessary to investigate deep learning techniques, such as Convolution Neural Network (CNN), Deep Belief Network (DBN), and Stacked Autoencoder (SAE), while obtaining various data with biosignal information (e.g., brain waves, pulse) and applying big data-based platform. Moreover, further studies will be conducted to analyze specific movement patterns of children by installing additional biometric data sensors, such as accelerometers and gyroscopes, to detect movements of children and establish a more precise stress-state identification model.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] Z. Karbasi, M. Kadivar, R. Safdari et al., "Better monitoring of abused children by designing a child abuse surveillance system: determining national child abuse minimum data set," *The International Journal of Health Planning and Management*, vol. 35, no. 4, pp. 843–851, 2020.

[2] J. S. Seo, "Through monitoring solution child abuse and children's rights at child care center," *Convergence Security Journal*, vol. 13, no. 2, pp. 55–66, 2013.

[3] C. H. Ryu and H. S. Yoon, "A study on the realities and preventive countermeasures of child abuse committed by biological parents by information analysis," *Journal of the Korea Society of Computer and Information*, vol. 24, no. 5, pp. 171–177, 2019.

[4] T.-Y. Kim, H. Ko, and S.-H. Kim, "Data analysis for emotion classification based on bio-information in self-driving vehicles," *Journal of Advanced Transportation*, vol. 2020, Article ID 8167295, 11 pages, 2020.

[5] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? Detection of stress patterns from GSR sensor data," in *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, pp. 573–580, Vancouver, BC, Canada, December 2011.

[6] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, 2005.

[7] Y. Jung and Y. I. Yoon, "Monitoring senior wellness status using multimodal biosensors," in *Proceedings of the 2016 International Conference on Big Data and Smart Computing (BigComp)*, pp. 435–438, IEEE, Hong Kong, China, January 2016.

[8] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410–417, 2010.

[9] P. Melillo, M. Bracale, and L. Pecchia, "Nonlinear heart rate variability features for real-life stress detection. Case study: students under stress due to university examination," *Biomedical Engineering Online*, vol. 10, no. 1, pp. 96–13, 2011.

[10] H. Kurniawan, A. V. Maslov, and M. Pechenizkiy, "Stress detection from speech and galvanic skin response signals," in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pp. 209–214, Porto, Portugal, 2013.

[11] P. Ruvolo and J. Movellan, "Automatic cry detection in early childhood education settings," in *Proceedings of the 2008 7th IEEE International Conference on Development and Learning*, pp. 204–208, Monterey, CA, USA, August 2008.

[12] R. M. Birn, B. J. Roeber, and S. D. Pollak, "Early childhood stress exposure, reward pathways, and adult decision making," *Proceedings of the National Academy of Sciences*, vol. 114, no. 51, pp. 13549–13554, 2017.

[13] L. Zhang, L. Jiang, and C. Li, "A new feature selection approach to naive Bayes text classifiers," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 30, no. 2, Article ID 1650003, 2016.

[14] S. Al-Eidi, O. Darwish, Y. Chen, and G. Husari, "SnapCatch: automatic detection of covert timing channels using image processing and machine learning," *IEEE Access*, vol. 9, p. 177, 2021.

[15] T.-Y. Kim, S.-H. Bae, and Y.-E. An, "Design of smart home implementation within IoT natural language interface," *IEEE Access*, vol. 8, pp. 84929–84949, 2020.

[16] J.-H. Roh, H.-J. Kim, and J.-Y. Chang, "Improving hypertext classification systems through WordNet-based feature abstraction," *The Journal of Society for E-Business Studies*, vol. 18, no. 2, pp. 95–110, 2013.

[17] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: a survey," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 3, pp. 1287–1301, 2012.

[18] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, and M. Griss, "Activity-aware mental stress detection using physiological sensors," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, Berlin, Germany, pp. 282–301, 2012.

[19] K. S. Alishamol, T. T. Fousiya, K. J. Babu, M. Sooryadas, and L. Mary, "System for infant cry emotion recognition using DNN," in *Proceedings of the 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 867–872, Tirunelveli, India, August 2020.

[20] L. Abou-Abbas, H. Fersaie Alaie, and C. Tadj, "Automatic detection of the expiratory and inspiratory phases in newborn cry signals," *Biomedical Signal Processing and Control*, vol. 19, pp. 35–43, 2015.

[21] A. Rosales-Pérez, C. A. Reyes-García, J. A. Gonzalez, O. F. Reyes-Galaviz, H. J. Escalante, and S. Orlandi, "Classifying infant cry patterns by the genetic selection of a fuzzy model," *Biomedical Signal Processing and Control*, vol. 17, pp. 38–46, 2015.

[22] F. Shahbazi and B. M. Asl, "Generalized discriminant analysis for congestive heart failure risk assessment based on long-term heart rate variability," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 2, pp. 191–198, 2015.

[23] A. Eckstein and P. P. Vlachos, "Assessment of advanced windowing techniques for digital particle image velocimetry (DPIV)," *Measurement Science and Technology*, vol. 20, no. 7, Article ID 075402, 2009.

[24] C. McKay, I. Fujinaga, and P. Depalle, "jAudio: a feature extraction library," in *Proceedings of the International Conference on Music Information Retrieval*, pp. 600–603, London, UK, September 2005.

[25] C. McKay, *Automatic Music Classification with jMIR*, McGill University, Montreal, Canada, 2010.

[26] H. Zhang, B. Wen, J. Liu, and Y. Zeng, "The prediction and error correction of physiological sign during exercise using bayesian combined predictor and naive Bayesian classifier," *IEEE Systems Journal*, vol. 13, no. 4, pp. 4410–4420, 2019.

[27] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[28] S. Kim, "Method and case study of decision tree for content design education," *The Journal of the Convergence on Culture Technology*, vol. 5, no. 4, pp. 283–288, 2019.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory-COLT'92*, pp. 144–152, New York; NY, USA, 1992.

[31] U. Fayyad and K. Irani, *Multi-interval Discretization of Continuous-Valued Attributes for Classification Learning*, http://hdl.handle.net/2014/35171, 1993.

[32] Y. Fu, Z. Yin, M. Su, M. Wu, and G. Liu, "Construction and reasoning approach of Belief rule-base for classification base on decision tree," *IEEE Access*, vol. 8, pp. 138046–138057, 2020.

[33] S.-H. Lee and D.-W. Lee, "A study on u-health fusion field based on Internet of Thing," *Journal of the Korea Convergence Society*, vol. 7, no. 4, pp. 19–24, 2016.

[34] M. Collotta, G. Pau, T. Talty, and O. K. Tonguz, "Bluetooth 5: a concrete step forward toward the IoT," *IEEE Communications Magazine*, vol. 56, no. 7, pp. 125–131, 2018.

[35] C. Crema, A. Depari, A. Flammini et al., "Smartphone-based system for the monitoring of vital parameters and stress conditions of amatorial racecar drivers," in *Proceedings of the 2015 IEEE SENSORS*, pp. 1–4, IEEE, Busan, South Korea, 2015.

[36] Y. S. Can, N. Chalabianloo, D. Ekiz, J. Fernandez-Alvarez, G. Riva, and C. Ersoy, "Personal stress-level clustering and decision-level smoothing to enhance the performance of ambulatory stress detection with smartwatches," *IEEE Access*, vol. 8, pp. 38146–38163, 2020.

[37] AudioMicro, http://www.audiomicro.com.

[38] H. Hamout and A. Elyousfi, "Fast depth map intra coding for 3D video compression based tensor feature extraction and data analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 7, p. 1, 2019.

[39] Z. Liu, Y. Wang, and T. Chen, "Audio feature extraction and analysis for scene segmentation and classification," *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, vol. 20, no. 1, pp. 61–79, 1998.

[40] B. Barabasz, M. Łoś, M. Woźniak, L. Siwik, and S. Barrett, "Coupled isogeometric finite element method and hierarchical genetic strategy with balanced accuracy for solving optimization inverse problem," *Procedia Computer Science*, vol. 108, pp. 828–837, 2017.

[41] M. B. Gorzałczany and F. Rudziński, "A multi-objective genetic optimization for fast, fuzzy rule-based credit classification with balanced accuracy and interpretability," *Applied Soft Computing*, vol. 40, pp. 206–220, 2016.