

Article

Not peer-reviewed version

A Negative Emotion Recognition System with IoT-Based Multimodal Biosignal Data

[Seung-mi Ham](#) , Hye-min Lee , Jae-hyun Lim , [Jeongwook Seo](#) *

Posted Date: 22 September 2023

doi: 10.20944/preprints202309.1560.v1

Keywords: Emotion Recognition; Multimodal; Biosignal; Wearable Device; Internet of Things; Support Vector Machine




Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

A Negative Emotion Recognition System with IoT-Based Multimodal Biosignal Data

Seung-mi Ham ¹, Hye-min Lee ¹, Jae-hyun Lim ² and Jeongwook Seo ^{1,*} 

¹ School of Computing and Artificial Intelligence, Hanshin University, Osan, 18101 Korea

² Platform Development Team, Wright Brothers Co., Ltd., Seoul, 04384 Korea; windsor@wrightbrothers.kr

* Correspondence: jwseo@hs.ac.kr

Abstract: Previous studies to recognize negative emotions for mental health care have used heavy equipment directly attaching electroencephalogram (EEG) electrodes to the head, and they have proposed binary classification methods to identify negative emotions. To tackle this problem, we propose a negative emotion recognition system to collect multimodal biosignal data such as five EEG signals from an EEG headset and heart rate, galvanic skin response, and skin temperature from a smart band for classifying multiple negative emotions. It consists of an android Internet of Things (IoT) application, a oneM2M-compliant IoT server, and a machine learning server. The android IoT application uploads the biosignal data to the IoT server. By using the biosignal data stored in the IoT server, the machine learning server recognizes the negative emotions of disgust, fear, and sadness using a multi-class support vector machine (SVM) model with a radial basis function kernel (RBF). The experimental results demonstrate that the multimodal biosignal data approach achieves 93% accuracy. Moreover, when considering only data from the smart band, it could achieve 98% accuracy by optimizing the hyper-parameter. Based on these results, we plan to develop a system that accurately detects negative emotions in real-time and expresses those emotions in the metaverse.

Keywords: emotion recognition; multimodal; biosignal; wearable device; Internet of Things; support vector machine

1. Introduction

As the development of IT technology and interest in personal health increase, the demand for wearable devices such as smart bands and smartwatches is increasing [1,2]. These devices enable the collection of various biosignal data, and the collected biosignal data is applied in various fields such as health, medical care, and emotion recognition [3–5]. In addition, with the popularization of user-customized intelligent services and the development of artificial intelligence technology, research on emotion recognition using biosignal data is becoming more important [6–8]. Previous research on emotion recognition has developed voice recognition, facial expression recognition, and gesture recognition. However, these studies often face challenges in accurately identifying actual emotional states because individuals may intentionally hide or distort their true feelings [9]. On the other hand, biosignals have the characteristic of not being intentionally manipulated since they change without one's awareness [10]. Furthermore, biosignals can objectively recognize not only surface emotions but also inner emotional states. For this reason, emotion recognition studies using biosignal data have been attracting attention. They have usually focused on classifying various emotion types, so they recognized six basic emotions: happiness, sadness, anger, disgust, fear, and surprise [11]. However, to manage the negative emotions that harm our mental health in our daily lives, it is necessary to focus on negative emotions (e.g. disgust, fear, sadness). Previous studies for recognizing negative emotions have mainly used a unimodal approach using only a single type of sensor to determine negative emotions. In particular, they typically attached electroencephalogram (EEG) electrodes directly to the head or used complex heavy equipment systems [12–14]. But, this approach is difficult and complex for ordinary people to wear, and has limitations in everyday use. If they use a multimodal approach that effectively utilizes different types of sensor data instead of an unimodal approach, the accuracy

of the emotion recognition model can be further improved [15]. In addition, they proposed a binary classification method to determine whether or not there is predominantly negative emotion [16]. To tackle this problem, we chose disgust, fear, and sadness as specific labels for negative emotions [17–19]. Therefore, we propose a negative emotion recognition system that classifies multiple negative emotions (disgust, fear, sadness) by collecting multimodal biosignal data through the two easy-to-use wearable Internet of Things (IoT) devices. Our aim is to enhance the accuracy of classifying various negative emotions (disgust, fear, sadness) and make them more accessible to users in their daily lives. The system utilizes an EEG headset and a smart band to collect multimodal biosignal data, including five EEG signals (alpha waves, beta waves, gamma waves, delta waves, theta waves), heart rate (HR), galvanic skin response (GSR), and skin temperature (SKT). The multimodal biosignal data is continuously and conveniently collected. We build a multi-class support vector machine (SVM) model with a radial basis function kernel (RBF) by labeling emotions on the collected multimodal biosignal data. In building the SVM model, we compare and analyze the accuracy of using all wearable device data versus using only the smart band data. The purpose is to examine whether good performance can be achieved by using only some data. We compare the results of the two models to determine which approach worked better. In addition, we create an optimal model by finding the values of the parameters C and γ that are most suitable for the model.

2. Proposed a Negative Emotion Recognition System

Figure 1 shows the proposed negative emotion recognition system with IoT-based multimodal biosignal data to manage negative emotions in daily life.

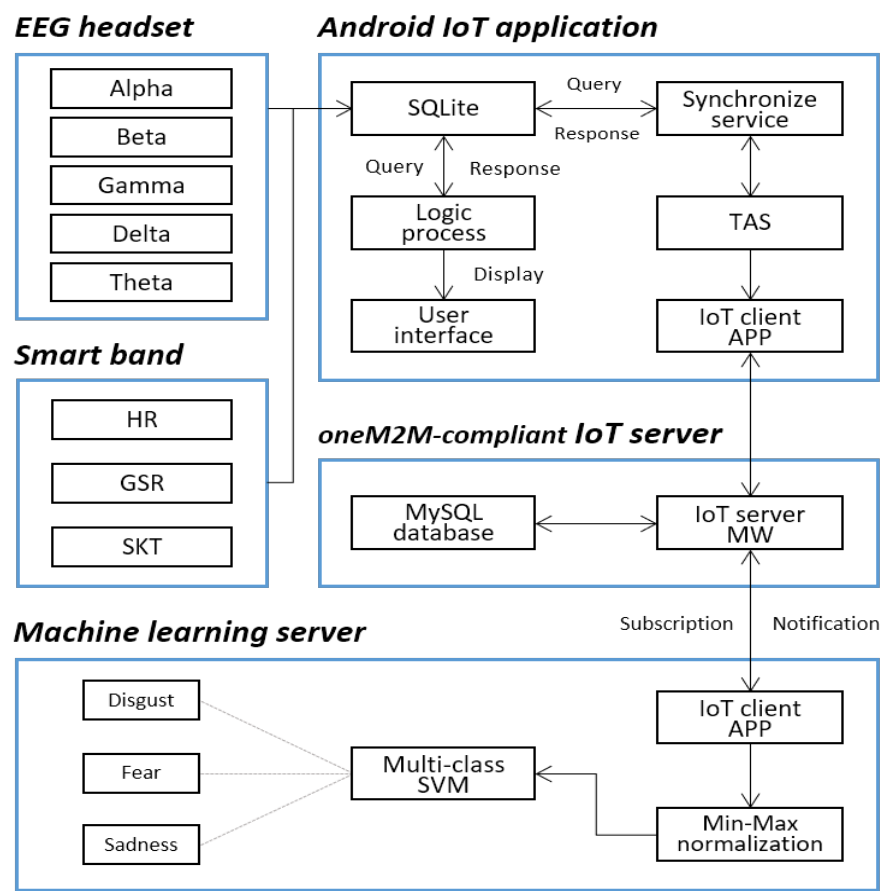


Figure 1. A negative emotion recognition system with IoT-based multimodal biosignal data.

Our proposed system consists of an android Internet of Things (IoT) application, a oneM2M-compliant IoT server, and a machine learning server. We will briefly explain the flow

of the system. The Android IoT application collects biosignal data and uploads it to the IoT server. The machine learning server receives biosignal data stored in the IoT server using the subscription and notification functions provided by oneM2M and recognizes negative emotions such as disgust, fear, and sadness.

First of all, We use an EEG headset and a smart band as two wearable IoT devices to collect multimodal biosignal data. The EEG headset is worn on the head to measure brainwave activity, including alpha, beta, gamma, delta, and theta waves. It uses a simple method of placing the measuring electrode sensor on the skin of the forehead and attaching the ear clip to the earlobe. The smart band is worn on the wrist and measures biosignals such as HR, GSR, and SKT through sensors built into the band. The two easy-to-use wearable IoT devices transmit the multimodal biosignal data to the android IoT application (APP) installed on the smartphone through Bluetooth wireless communication. The android IoT APP checks the received multimodal biosignal data change and then saves the data in the SQLite database. Synchronize service retrieves multimodal biosignal data from the SQLite database once a second and transmits it to TAS, and TAS transmits it to an IoT client APP using client socket communication.

The IoT client APP, modeled as a dedicated node application entity (ADN-AE), sends a request to the IoT server middleware (MW), modeled as an infrastructure node common service entity (IN-CSE), to register itself [20]. Here, ADN is a node that contains at least one AE and does not contain a common service entity (CSE), and AE is an entity in the application layer that implements application service logic for end-to-end IoT solutions. In addition, the IN is a node that contains one CSE and can also contain AEs, and a CSE is an entity implemented with 12 common service functions (CSFs) that can be used in common in various application entities of IoT. The 12 CSFs include registration, discovery, security, group management, and subscription & notification, etc [21,22]. When the IoT client APP receives a response that the ADN-AE is successfully registered, it requests the IoT server MW to create a container for the EEG headset and a container for the smart band. The IoT server MW creates two containers in the ADN-AE and then sends a response to the IoT client APP that the creation is complete. The IoT client APP creates and sends the content instances containing five EEG signals and the content instances containing HR, GSR, and SKT to the IoT server MW. Then, the IoT server MW checks the content instances sent by the IoT client APP and stacks the five EEG signals in the EEG headset container and the HR, GSR, and SKT in the container for the smart band.

The IoT client APP in the machine learning server subscribes to the resources of the IoT server MW. When a new content instance is uploaded to the IoT server MW, the IoT client APP receives the notification message sent by the IoT server MW and extracts the notification message to obtain the multimodal biosignal data. After that, it normalizes the multimodal biosignal data through a min-max scaler and feeds the normalized data into our multi-class SVM model with a non-linear RBF kernel to classify the negative emotions such as disgust, fear, and sadness.

3. Exploratory Data Analysis and Multi-class SVM Model

3.1. Exploratory Data Analysis

We collected the multimodal biosignal data of 30 participants' 7 emotions: happiness, amusement, disgust, neutral, fear, tenderness, and sadness. As mentioned before, we specifically chose disgust, fear, and sadness as labels to represent negative emotions among the seven emotions. Then, we focused on analyzing multimodal biosignal data corresponding to three specific negative emotions. We used exploratory data analysis techniques to identify the characteristics, patterns, and structural relationships of the collected multimodal biosignal data.

Figure 2 shows the histograms of the multimodal biosignal data associated with the negative emotions to observe their distributions within specific intervals, facilitating the identification of sections with higher frequencies. Additionally, it allows us to estimate their mean, median, and central tendency and detect any asymmetries and outliers. As a result of the analysis, the five EEG signals mainly show

a symmetrical distribution and have a single peak. On the other hand, HR, GSR, and SKT data show an asymmetric distribution and have two or more peaks. In particular, the graphs of HR and GSR display a right-skewed shape with the peak on the left and the tail extending to the right.

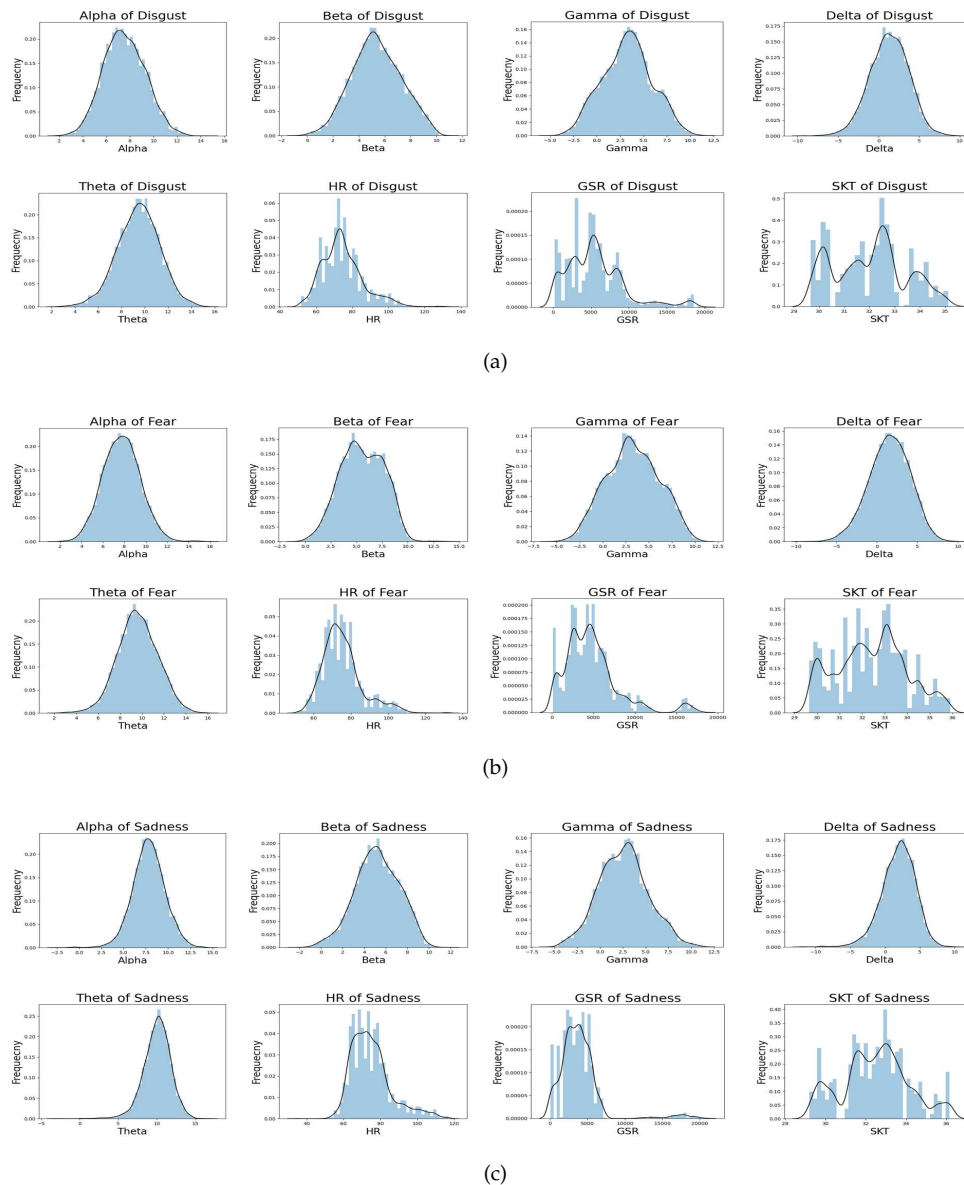


Figure 2. Histograms of the multimodal biosignal data measured from an EEG headset and a smart band according to three negative emotions: (a) Disgust. (b) Fear. (c) Sadness.

We calculate the covariance between two variables to see how they change together. Covariance has the problem of being affected by the units of the variable. However, the correlation coefficient is not affected by the units of the variables because it non-dimensionally expresses the linear relationship between two variables. In statistics, there are various methods to obtain correlation results between two variables, such as Kendall, Pearson, Spearman, etc. We expected that the given data would exhibit a linear relationship between variables, so we chose the Pearson correlation coefficient, which excels in measuring linear relationships between variables. On the other hand, Spearman correlation is suitable for ranked data, and Kendall's Tau is used to measure correlation while preserving the order of data. For these reasons, we determined that the Pearson correlation coefficient was the most suitable for our research objectives. In Figures 3–5, the Pearson correlation coefficient is used to characterize the

degree of linear correlation between two variables [23]. The Pearson correlation coefficient between two variables (x, y) is the quotient of the covariance and standard deviation between the two variables, which is defined as:

$$r(x, y) = \frac{\sum_{n=1}^N (x_n - \bar{x})(y_n - \bar{y})}{\sqrt{\sum_{n=1}^N (x_n - \bar{x})^2 \sum_{n=1}^N (y_n - \bar{y})^2}} \quad (1)$$

where N is the number of samples, and x and y represent the series to be analyzed. \bar{x} and \bar{y} represent the mean value of the observed series. The resulting value of $r(x, y)$ ranges from -1 to 1. A value of 1 indicates a positive correlation where all data points fall well on a straight line and increase as the value increases. A value of -1 means all data points fall on a straight line and a negative correlation decreases as the coefficient increases. A value of 0 means there is no linear relationship between the two variables [24,25].

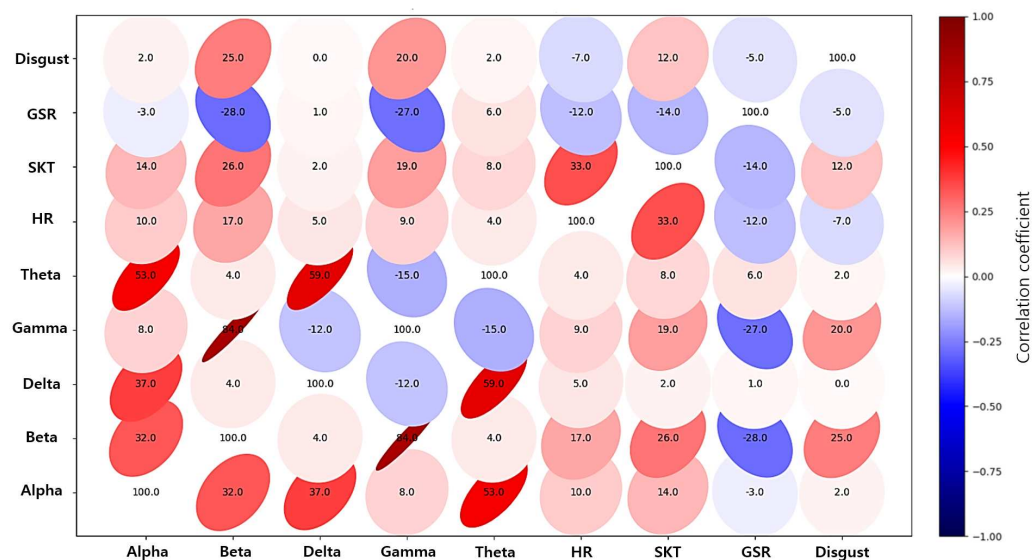


Figure 3. Correlation matrix of multimodal biosignal data for the disgust state.

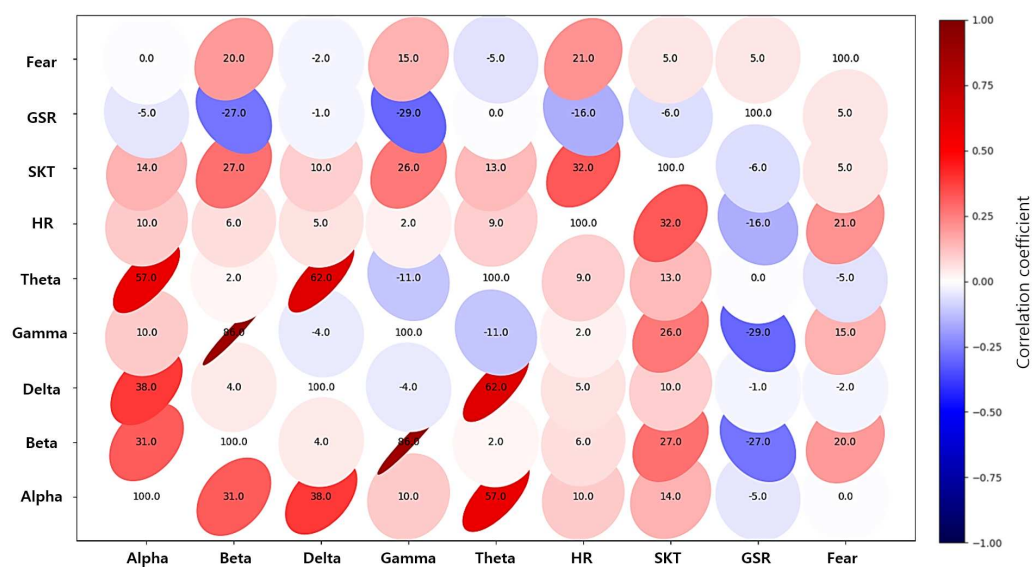


Figure 4. Correlation matrix of multimodal biosignal data for the fear state.

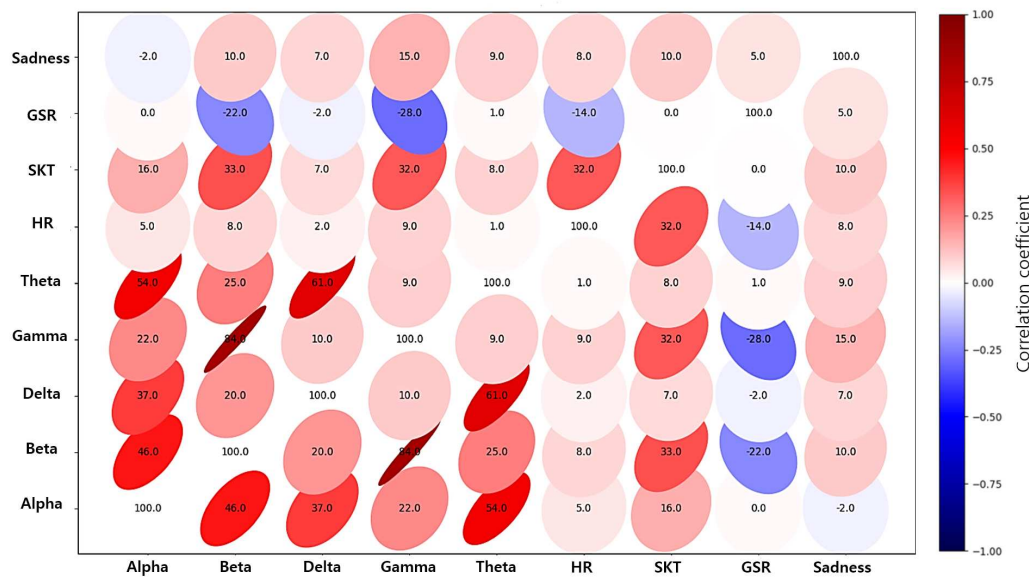


Figure 5. Correlation matrix of multimodal biosignal data for the sadness state.

The correlation matrices made from the Pearson correlation coefficient are shown where the larger the absolute value of the coefficient, the narrower the ellipse, and the smaller the absolute value, the wider the ellipse appears. As a result of the analysis, in the disgust state, beta waves and gamma waves showed positive values of 25% and 20%, respectively. These findings suggest a strong positive correlation between beta and gamma waves in the state of disgust. In the fear state, beta waves, gamma waves, and HR showed positive values of 20%, 15%, and 21%, respectively. This indicates a significant positive correlation between the fear state and beta waves, gamma waves, and HR. Finally, in the sadness state, beta waves, gamma waves, and SKT showed positive values of 10%, 15%, and 10%, respectively. These results suggest a strong positive correlation between the state of sadness and beta waves, gamma waves, and SKT.

3.2. Multi-class Non-linear SVM Model

Machine learning models such as decision tree, random forest, k-nearest neighbor (KNN), artificial neural network (ANN), and SVM are frequently used to classify data. The decision tree is inflexible in parametric modeling, random forest is limited to handling high-dimensional data sets, and KNN is quite slow and very sensitive to outliers when dealing with large amounts of data. ANN makes their internal parameters difficult to adjust, which can lead to underfitting and overfitting problems [26]. On the other hand, SVM has high classification accuracy and a good ability to handle high-dimensional and large data sets [27]. Therefore, it is widely used in the field of classification and is suitable for classifying linear and non-linear data. It also achieves high accuracy by using kernel functions that map the data into higher dimensions, improving the classification of non-linear data [28]. In this study, considering the need for a technique specialized in handling non-linear elements, we chose a non-linear SVM model. Specifically, we chose the RBF kernel, known for its effectiveness in handling datasets with non-linear data distributions. The RBF kernel facilitates linear separation in high-dimensional spaces by virtue of non-linear feature mapping and proves highly valuable when confronted with diverse types of data. Consequently, we propose a multiclass SVM with the RBF kernel. The multimodal biosignal dataset for training the proposed model is divided into an 80% training set and a 20% test set. We normalize all features of training and test data by scaling them within the range [0, 1] through Min-Max scaler to make it more suitable for training. A multimodal biosignal data set labeled with a training pattern can be represented by

$$D_N = \{(x_1, y_1), \dots, (x_N, y_N)\} = \{(x_n, y_n)\}_{n=1}^N \quad (2)$$

where $\mathbf{x}_n = [x_{1,n}, x_{2,n}, \dots, x_{K,n}]^T$ denotes a training feature vector of length K , and $y_n \in \{1, 2, \dots, M\}$ denotes a label among M classes. The SVM model finds the marginally separated hyperplane that maximally represents the distance from each class's hyperplane to the nearest feature vector, called the support vector. It utilizes a kernel function to map data from a low-dimensional space to a high-dimensional space to create a non-linearly separable pattern. A multi-class SVM model built using an RBF kernel is represented by Equation (3). It splits a multi-class data set into multiple binary classification problems and uses a one-versus-rest method, a heuristic method using binary classifiers.

$$\hat{y}_{\text{new}} = \underset{m=1,2,\dots,M}{\operatorname{argmax}} \sum_{n \in S} \alpha_{m,n} y_{m,n} L(\mathbf{x}_n, \mathbf{x}_{\text{new}}) + b_m \quad (3)$$

where S denotes the set of support vectors, and $\alpha_{m,n}$ represents the Lagrange multiplier with the constraint $0 \leq \alpha_{m,n} \leq C$ on class m . C denotes the hyper-parameter that determines the trade-off for margin and error, $y_{m,n}$ represents the temporary label, and b_m denotes the bias term. In addition, the RBF kernel is represented by Equation (4).

$$L(\mathbf{x}_n, \mathbf{x}_{\text{new}}) = \exp(-\delta \|\mathbf{x}_n - \mathbf{x}_{\text{new}}\|_2^2) \quad (4)$$

where δ denotes a hyper-parameter that defines the extent to which the influence of training vector \mathbf{x}_n reaches the support vectors, and $\|\cdot\|_2$ denotes the L2 norm [29]. The smaller the value of δ , the greater the effect of \mathbf{x}_n on the classification of \mathbf{x} , resulting in a larger variance. The larger the value of δ , the smaller the effect of \mathbf{x}_n on the classification of \mathbf{x} , resulting in a smaller variance.

4. Experimental results

4.1. Experimental Conditions

We use an EEG headset and a smart band, which are commercially available wearable IoT devices, to collect various biosignal data of alpha waves, beta waves, gamma waves, delta waves, theta waves, HR, GSR, and SKT. For our experiments, we used emotional video clips containing seven emotion categories covering common emotional contexts: anger, sadness, fear, disgust, amusement, tenderness, and neutral [30]. The video clips used as emotional stimuli were produced in English and consisted of 10 specific movie scenes most cited as evoking each of the seven emotions. [30] states that the clip successfully met various validity criteria and confirmed its validity through a multidimensional evaluation experiment. Therefore, it is mentioned that it consists of movies that are most likely to elicit related emotions most effectively. We collected multimodal biosignal data by asking the experimenter to watch a video while wearing wearable devices, to capture various emotions. And then, we generated emotion labels accordingly.

Figure 6 illustrates the experimental components of the proposed emotion recognition system. The components of the experiment are for collecting multimodal biosignals in various emotions. It includes an EEG headset and smart band belonging to two wearable IoT devices, an android IoT APP, a oneM2M-compliant IoT server, and movie clips based on emotions. Figure 7 shows an experimental process to collect multimodal biosignal data of participants through watching movie clips. In the experiment, a total of 30 healthy young subjects participated, comprising 23 males and 7 females, with a mean age of 23.2 ± 2.1 years (20-27). Although gender imbalances exist in our study, we found in [30] that differences in subjective arousal levels and negative affect between women and men did not reach statistical significance. We scaled the video clips that participants would watch to the same resolution (1920 × 1080) and displayed them on a 24-inch LCD screen. Additionally, we used two speakers to watch and adjusted the volume to an appropriate level so that participants could watch comfortably [31]. Participants wear an EEG headset on their heads and a smart band on their wrists. We randomly played the movie clips to collect data on participants' various emotions and prevented continuous playback of movie clips with similar emotions. While watching the video,

the participants filled out the prepared questionnaire. Figure 8 shows a questionnaire designed to self-record participants' seven emotions while they watch movie clips. We adopted several approaches to consider the subjective nature and various individual differences in emotional responses. In the survey, we collected and analyzed participants' characteristics, such as gender, age, and family composition. Consequently, participants recorded their personal information in the questionnaire, viewed movie clips designed to elicit seven emotions, and provided responses regarding the intensity of each emotion they experienced.

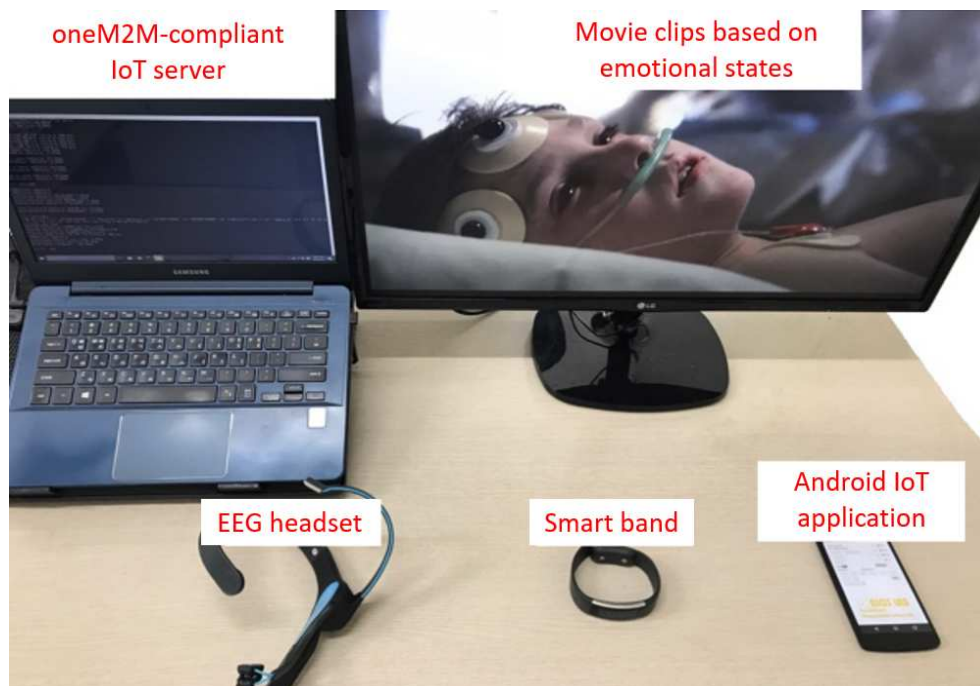


Figure 6. Experimental components for multimodal biosignals collection in various emotions.

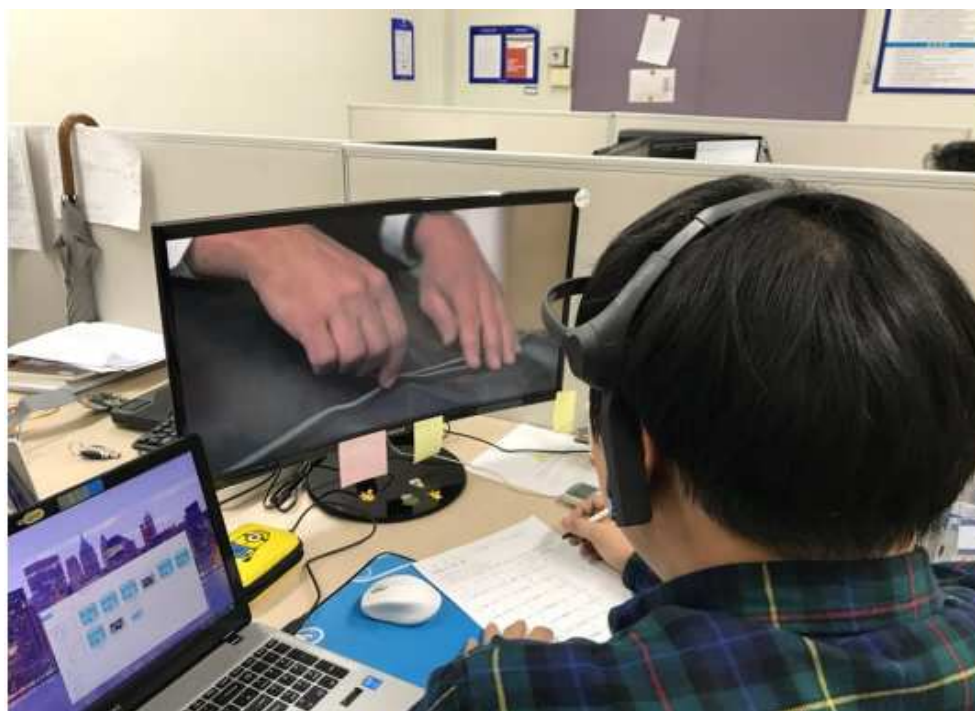


Figure 7. Experimental process for collecting multimodal biosignals through watching movie clips.

Name :	Sex : male() female()	Student ID :		
Emotion recognition video evaluation				
Thank you for participating in the survey. This survey evaluates your emotions after watching each video. This data will only be used for the research of the paper. Please indicate your satisfaction with each question. (1 = low, 2 = slightly low, 3 = normal, 4 = slightly high, 5 = high)				
Age :	Family background :			
Video start time :	Video end time :			
How angry were you when you saw the first video?				
1	2	3	4	5
How sad were you when you saw the second video?				
1	2	3	4	5
How fearful were you when you saw the third video?				
1	2	3	4	5
How disgusted were you when you saw the fourth video?				
1	2	3	4	5
How amused were you when you saw the fifth video?				
1	2	3	4	5
How tender did you feel when you saw the sixth video?				
1	2	3	4	5
How neutral did you feel when you saw the seventh video?				
1	2	3	4	5

Figure 8. A questionnaire to record seven emotions while watching a video.

4.2. Performance Evaluation

Figure 9 shows the experimental results of linear SVM model and non-linear SVM model with an RBF kernel (default), and non-linear SVM using an RBF kernel with specific parameters ($C = 10$, $\gamma = 10$). In Figure 9, F1 to F8 are cases in which all multimodal biosignal data (five EEG signals, HR, GSR, and SKT) collected from two wearable IoT devices are used, and F6 to F8 are cases in which only HR, GSR, and SKT collected from the smart band are used. Through this, we compared the average accuracy of the linear SVM model and the non-linear multi-class SVM and confirmed that the multi-class SVM model achieved higher accuracy. When comparing the linear SVM model and the multi-class SVM (default) using the RBF kernel, F1 to F8 are more accurate than F6 to F8, but the difference in accuracy is not as large as about 2.7 to 2.8. This suggests that building a model with only HR, GSR, and SKT collected from a smart band may be more efficient than using all of the multimodal biosignal data collected from two wearable IoT devices. Also, through a comparison of multi-class SVM using an RBF kernel (default) and multi-class SVM using an RBF kernel ($C = 10$, $\gamma = 10$), it can be confirmed that parameter values have a significant effect on accuracy. Therefore, it indicates that higher accuracy can be achieved by selecting appropriate parameter values and incorporating them into the model.

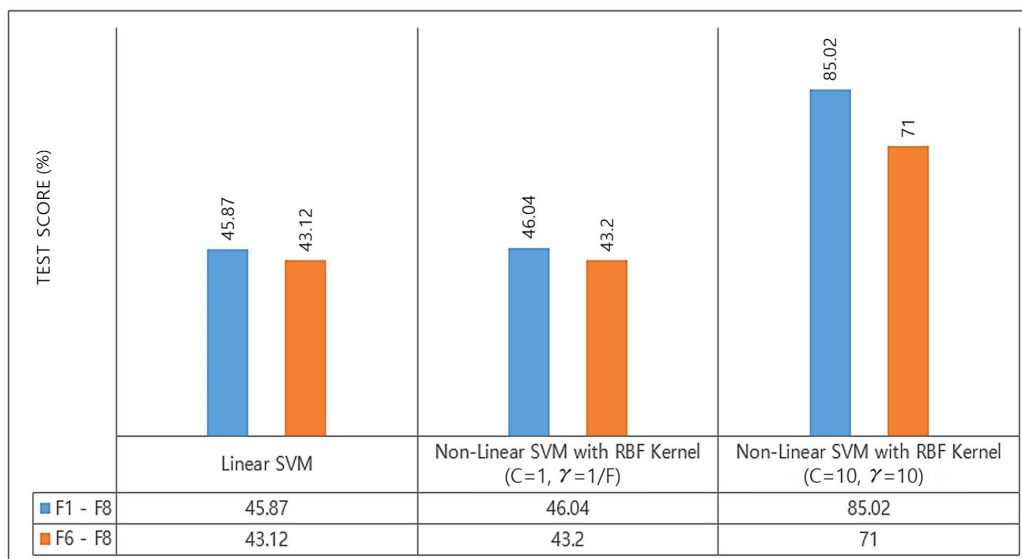


Figure 9. Average accuracy comparison according to SVM model.

A grid search was performed to determine the optimal parameter values for the model, and the accuracy according to parameter adjustment was shown as a heatmap as shown in Figures 10 and 11. When performing a grid search, the parameter values for C and γ ranged from 10^{-3} to 10^3 . Figure 10 shows a heatmap representing the accuracy for different parameter values of the multi-class SVM (with an RBF kernel) consisting of multimodal biosignal data (five EEG signals, HR, GSR, and SKT). The highest accuracy of 93% was achieved with parameter values of $C = 10, 10^2, 10^3$, and $\gamma = 10^2$. Also, Cross-validation confirmed that the most appropriate model was obtained when $C = 10$ and $\gamma = 10^3$. Figure 11 displays a heatmap showing the accuracy according to the parameter values of the multi-class SVM (with an RBF kernel) consisting of HR, GSR, and SKT. It can be seen that the highest accuracy of 98% was achieved when the parameter values are $C = 10, 10^2, 10^3$, and $\gamma = 10^3$. Moreover, cross-validation has shown that the optimal model was obtained when $C = 10^3$ and $\gamma = 10^3$. Therefore, it demonstrates that by adjusting the parameter values, the model consisting of HR, GSR, and SKT collected from a smart band can achieve higher accuracy than the model consisting of multimodal biosignal data collected from two wearable IoT devices.

Table 2 shows the performance of a multiclass SVM (with an RBF kernel) consisting of HR, GSR, and SKT from the smart band. The evaluation indicators used for performance comparison are introduced and defined as shown in Table 1. Following the above definition, three widely used metrics: Precision, Recall, and F1 score, can be calculated as follows.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2(Precision * Recall)}{Precision + Recall} \quad (7)$$

Precision refers to the model's ability to distinguish between negative emotions. Higher precision values improve the model's ability to distinguish between similar negative emotions. Recall reflects the proportion of models capable of solving the intraclass similarity problem. A higher recall value improves the model's ability to detect negative emotions with intra-class similarity. The F1 score

represents the balance between precision and recall. The higher the F1 score, the more powerful the model performs.

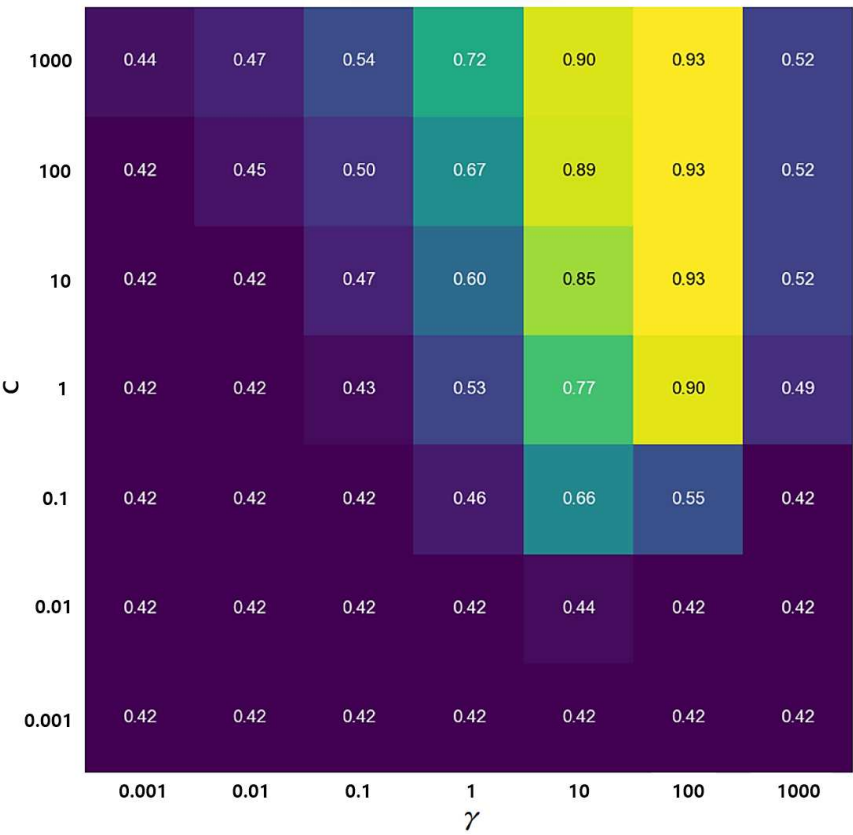


Figure 10. Heatmap according to multimodal biosignal data of two wearable IoT devices.

Table 1. Definition of terms for evaluation indicators.

True positive (TP)	A target traffic sign has been predicted to a correct type.
True negative (TN)	A non-target traffic sign has been predicted to a correct type.
False positive (FP)	A target traffic sign has been predicted to a wrong type.
False negative (FN)	A non-target traffic sign has been predicted to a wrong type.
Macro average	Averaging the unweighted mean per label
Weighted average	Averaging the support-weighted mean per label

Table 2. Performance of multiclass SVM consisting of HR, GSR, and SKT.

	Precision	Recall	F1 score	Support
Disgust	0.97	0.97	0.97	1504
Fear	0.98	0.98	0.98	2395
Sadness	0.99	0.99	0.99	1707
Accuracy			0.98	5606
Macro average	0.98	0.98	0.98	5606
Weighted average	0.98	0.98	0.98	5606

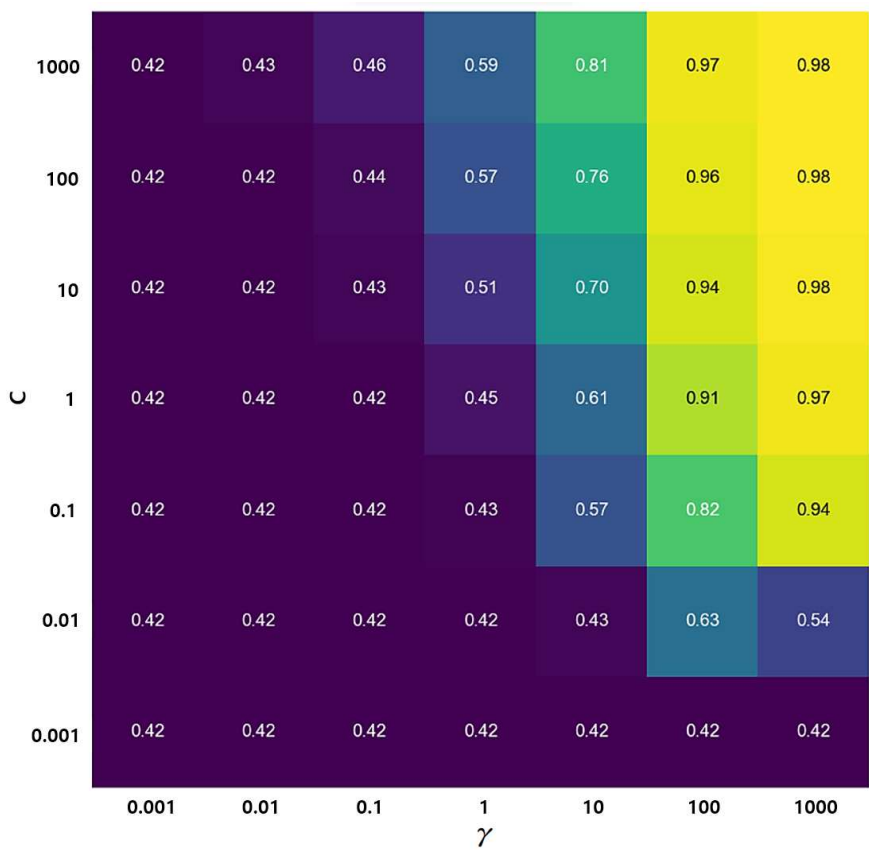


Figure 11. Heatmap according to biosignal data of a smart band.

Figure 12 illustrates the learning curve of a multi-class SVM model ($C = 10^3, \gamma = 10^3$) constructed using HR, GSR, and SKT data from the smart band. The model’s score is expressed as a percentage and shows how accuracy changes as the amount of training data increases. It presents the model’s performance on both the training and cross-validation sets. In the graph, the training score decreases from 99.8% to 99.2%, and the cross-validation score increases from 92.7% to 98.1%. These results indicate that both the training curve and cross-validation curve of the model perform well, confirming the absence of an overfitting problem. Additionally, we have verified the appropriateness of the parameter values (C and γ) selected for the model.

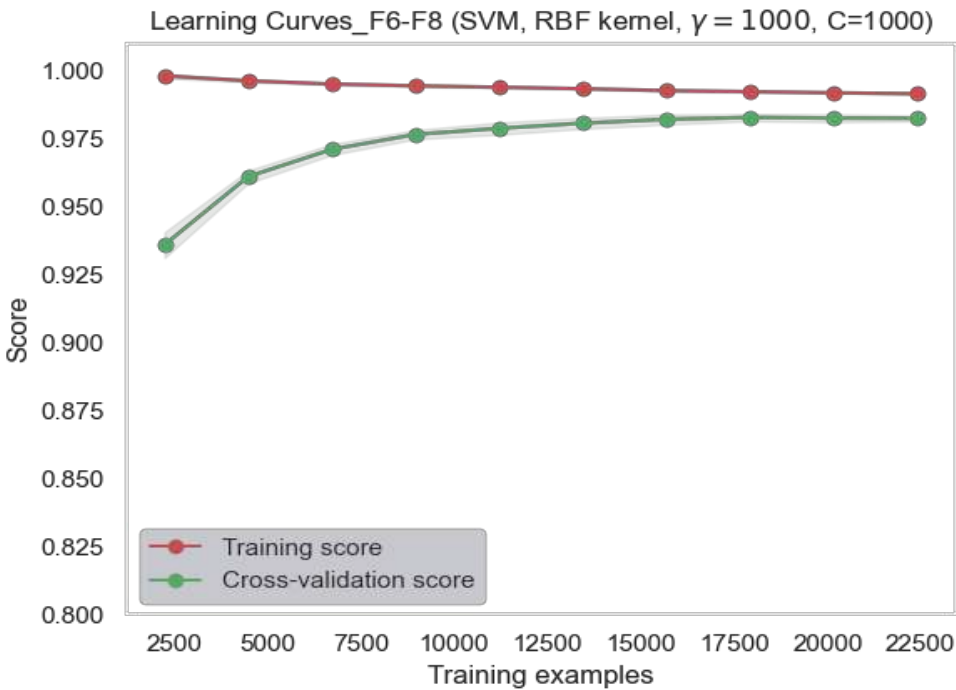


Figure 12. Learning curve for multi-class SVM model built using HR, GSR and SKT.

5. Discussion

We compare our study with studies that have investigated negative emotion recognition using biometric data, similar to our approach. Table 3 summarizes the comparison. Key observations are summarized as follows:

Table 3. Comparisons between our work and existing works.

Study	Year	Emotion	Biosignal	Data Collection Tool	Method
Ihmig et al. [32]	2020	Anxiety	ECG, GSR, RSP	Biofeedback system, BITalino biosignal measurement device, Ag/AgCl electrodes	Bagged trees
Lee and Yoo [12]	2020	Negative	ECG, GSR, SKT	MP 150TM of BIOPAC	LSTM
Al-Jumaily et al. [33]	2021	Stress	ECG, EMG	Certain sensors	Gaussian kernel SVM
Mekruksavanich et al. [34]	2022	Tension	ECG, EMG, GSR	Chest-worn equipment	CNN, ResNeXt
Our study	2023	Disgust, fear, sadness	HR, GSR, SKT	Microsoft Band 2	RBF kernal SVM

In [32], they conducted a study aimed at classifying anxiety levels into two and three categories by utilizing a combination of features extracted from ECG (electrocardiogram), GSR, and RSP (respiration) signals. They used a biofeedback system, BITalino biosignal measurement device, and Ag/AgCl electrode to collect biosignal data, and used Bagged Tree as a model learning method. In [12], they proposed a model that recognizes neutral and negative emotions by utilizing biosignals of ECG, GSR, and SKT. They collected biosignal data using the MP 150TM of BIOPAC Systems and applied

long short-term memory (LSTM) to design an emotion recognition model. In [33], they proposed a Gaussian kernel SVM model to detect stress states. They collected biosignal data using specific sensors that measure ECG and electromyogram (EMG) and classified stress states. In [34], they proposed a framework to classify tension using biosignal data (ECG, EMG, GSR) from chest-worn equipment. They presented a convolutional neural network (CNN) based on ResNeXt [35].

Existing studies mainly adopt a data collection tool by having users wear heavy equipment, but this method has several limitations. First, these devices are large and heavy, which can cause discomfort to the user. Also, it is difficult to wear these devices for long periods of time, which limits movement and can limit data collection in real-world environments. Lastly, if users are required to attach certain sensors themselves, they must understand exactly how to use the equipment, which can be difficult for some users. On the other hand, considering user convenience, our study used the easily wearable Microsoft Band 2 to easily and conveniently collect HR, GSR, and SKT data. Furthermore, while previous studies classified negative emotions into stages or binary categories, our study proposed a more specific classification of negative emotions into disgust, fear, and sadness.

6. Conclusions

Existing research on recognizing negative emotions faces challenges related to user accessibility in their everyday lives and is limited to simplistic binary classification. Therefore, we proposed a negative emotion recognition system that classifies multiple negative emotions (disgust, fear, sadness) by collecting multimodal biosignal data through an EEG headset and a smart band. It is based on an IoT server to continuously and conveniently collect multimodal biosignal data from two easy-to-use wearable IoT devices: an EEG headset and a smart band. We confirmed that the collected data is continuously stored in real-time within the database of the oneM2M-based IoT server MW. Also, we more accurately classified the negative emotions (disgust, fear, and sadness) by constructing a multi-class SVM with an RBF kernel. By comparing the accuracy of linear SVM model and non-linear SVMs, we found that non-linear SVM models with an RBF kernel achieve higher accuracy. Also, we observed slightly higher accuracy when using all multimodal biosignal data collected from two wearable IoT devices in the graph of a linear SVM and a non-linear SVM with an RBF kernel (default) than when using only HR, GSR, and SKT. However, the difference in their accuracy was not significant. Through this, we confirmed that we can improve the efficiency of the model by utilizing only HR, GSR, and SKT. We found the optimal parameters $C = 10^3$ and $\gamma = 10^3$ for the model through grid search. In addition, we verified that the multi-class SVM model built using only HR, GSR, and SKT with optimal parameters achieved an average accuracy of 98%. This accuracy is 5% higher than the accuracy achieved by a model constructed using all collected multimodal biosignal data from the two wearable IoT devices. We plan to conduct future research based on the results of achieving high accuracy using only biosignal data from a smart band with optimal parameters. Our goal is to develop a system that accurately detects the negative emotions of users wearing smart bands in real time and injects and expresses those emotions in real time to digital humans within the metaverse. Furthermore, since our research primarily focuses on recognizing three specific negative emotions (disgust, fear, and sadness), there is a need to conduct further research to recognize a broader range of emotions or more subtle emotional states.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-2021-0-01816) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

References

1. Choi, K.H.; Yang, E.S. A study on the trend of healthcare device technology by biometric signal. *J Korea Entertain Ind Assoc.* **2020**, *14*, 165–176.
2. Ferreira, J.J.; Fernandes, C.I.; Rammal, H.G.; Veiga, P.M. Wearable technology and consumer interaction: A systematic review and research agenda. *Comput. Hum. Behav.* **2021**, *118*, 106710.

3. Mahloko, L.; Adebesein, F. A systematic literature review of the factors that influence the accuracy of consumer wearable health device data. In Proceedings of the e-Business, e-Services and e-Society, Skukuza, South Africa, 6–8 April 2020; 96–107.
4. Singh, B.; Zopf, E.M.; Howden, E.J. Effect and feasibility of wearable physical activity trackers and pedometers for increasing physical activity and improving health outcomes in cancer survivors: A systematic review and meta-analysis. *J Sport Health Sci.* **2022**, *11*, 184–193.
5. Domínguez-Jiménez, J.A.; Campo-Landines, K.C.; Martínez-Santos, J.C.; Delahoz, E.J.; Contreras-Ortiz, S. H. A machine learning model for emotion recognition from physiological signals. *Biomed Signal Process Control* **2020**, *55*, 101646.
6. Kim, G.; Choi, I.; Li, Q.; Kim, J. A CNN-based advertisement recommendation through real-time user face recognition. *Appl. Sci.* **2021**, *11*, 9705.
7. Zhang, H. Expression-EEG based collaborative multimodal emotion recognition using deep autoencoder. *IEEE Access* **2020**, *8*, 164130–164143.
8. Gu, X.; Cai, W.; Gao, M.; Jiang, Y.; Ning, X.; Qian, P. Multi-source domain transfer discriminative dictionary learning modeling for electroencephalogram-based emotion recognition. *IEEE Trans. Comput. Soc. Syst.* **2022**, *9*, 1604–1612.
9. Liu, Y.; Fu, G. Emotion recognition by deeply learned multi-channel textual and EEG features. *Future Generation Computer Systems* **2021**, *119*, 1–6.
10. Domínguez-Jiménez, J. A.; Campo-Landines, K. C.; Martínez-Santos, J. C.; Delahoz, E. J.; Contreras-Ortiz, S. H. A machine learning model for emotion recognition from physiological signals. *Biomedical signal processing and control* **2020**, *55*, 101646.
11. Liu, Y.; Ding, Y.; Li, C.; Cheng, J.; Song, R.; Wan, F.; Chen, X. Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network. *Comput. Biol. Med.* **2020**, *123*, 103927.
12. Lee, J.; Yoo, S.K. Recognition of negative emotion using long short-term memory with bio-signal feature compression. *Sensors* **2020**, *20*, 573.
13. Long, F.; Zhao, S.; Wei, X.; Ng, S.C.; et al. Positive and negative emotion classification based on multi-channel. *Front. Behav. Neurosci.* **2021**, *15*, 720451.
14. Suhaimi, N.S.; Mountstephens, J.; Teo, J. EEG-based emotion recognition: A state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* **2020**, 2020.
15. Jo, C.; Jung, H. Multimodal Emotion Recognition System using Face Images and Multidimensional Emotion-based Text. *The Journal of Korean Institute of Information Technology* **2023**, *21*, 39–47.
16. Ancillon, L.; Elgendi, M.; Menon, C. Machine Learning for Anxiety Detection Using Biosignals: A Review. *Diagnostics* **2022**, *12*, 1794.
17. She, W.; Lv, Z.; Tao, J.; Niu, M. Micro-expression recognition based on multiple aggregation networks. In 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Auckland, New Zealand, 7–10 December 2020;
18. Chatterjee, S. Drivers of helpfulness of online hotel reviews: A sentiment and emotion mining approach. *International Journal of Hospitality Management* **2020**, *85*, 102356.
19. Li, J.; Hu, R.; Mukherjee, M. Discriminative Region Transfer Network for Cross-Database Micro-Expression Recognition. In ICC 2022-IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022;
20. Lee, S.; Kil, W.; Roh, B.H.; Kim, S.J.; Kang, J. S. Novel Architecture of OneM2M-Based Convergence Platform for Mixed Reality and IoT. *CMC* **2022**, *71*, 51.
21. Lee, J.; Gilani, K.; Khatoon, N.; Jeong, S.; Song, J. SEIF: A Semantic-enabled IoT Service Framework for Realizing Interoperable Data and Knowledge Retrieval. *IEIE SPC* **2023**, *12*, 9–22.
22. Mante, S.; Vaddhiparthi, S.S.S.; Ruthwik, M.; Gangadharan, D.; Hussain, A.M.; Vattam, A. A Multi Layer Data Platform Architecture for Smart Cities using oneM2M and IUDX. In 2022 IEEE 8th WF-IoT **2022**, 1–6.
23. Peng, L.; Zheng, S.; Zhong, Q.; Chai, X.; Lin, J. A novel bagged tree ensemble regression method with multiple correlation coefficients to predict the train body vibrations using rail inspection data. *Mechanical Systems and Signal Processing* **2023**, *182*, 109543.
24. Šverko, Z.; Vrankić, M.; Vlahinić, S.; Rogelj, P. Complex Pearson Correlation Coefficient for EEG Connectivity Analysis. *Sensors* **2022**, *22*, 1477.

25. Ning, Z.; Wang, B.; Li, S.; Jia, X.; Xie, S.; Zheng, J. Pipeline risk factors analysis using the Pearson correlation coefficient method and the random forest importance factor method. In 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech), Split and Bol (Island of Brac), Croatia, 20–23 June 2023; pp. 1–5.
26. Otchere, D.A.; Ganat, T.O.A.; Gholami, R.; Ridha, S. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering* **2021**, *200*, 108182.
27. Choi, J.; Im, S. Leak Detection and Classification of Water Pipeline based on SVM using Leakage Noise Magnitude Spectrum. *Journal of The Institute of Electronics and Information Engineers* **2023**, *60*, 6–14.
28. Mohammadi, M.; Rashid, T.A.; Karim, S.H.T.; Aldalwie, A.H.M.; Tho, Q.T.; Bidaki, M.; Hosseinzadeh, M. A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. *Journal of Network and Computer Applications* **2021**, *178*, 102983.
29. Jebli, I.; Belouadha, F.Z.; Kabbaj, M.I.; Tilioua, A. Prediction of solar energy guided by pearson correlation using machine learning. *Energy* **2021**, *224*, 120109.
30. Schaefer, A.; Nils, F.; Sanchez, X.; Philippot, P. Assessing the effectiveness of a large database of emotion-eliciting films : a new tool for emotion researchers. *Cogn. Emot.* **2010**, *24*, 1153–1172.
31. Liu, Y.J.; Yu, M.; Zhao, G.; Song, J.; Ge, Y.; Shi, Y. Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* **2017**, *9*, 550–562.
32. Ihmig, F.R.; Neurohr-Parakenings, F.; Schäfer, S.K.; Lass-Hennemann, J.; Michael, T. On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals. *Plos one* **2020**, *15*, e0231517.
33. Al-Jumaily, A.A.; Matin, N.; Hoshayr, A.N. Machine learning based biosignals mental stress detection. In Soft Computing in Data Science: 6th International Conference, SCDS 2021, Virtual Event, 2–3 November 2021; 28–41.
34. Mekruksavanich, S.; Hnoohom, N.; Jitpattanakul, A. A Deep Residual-based Model on Multi-Branch Aggregation for Stress and Emotion Recognition through Biosignals. In 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Huahin, Thailand, 24–27 May 2022; 1–4.
35. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; 1492–1500.

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