

Research Article

A Study on User Recognition Using 2D ECG Image Based on Ensemble Networks for Intelligent Vehicles

Min-Gu Kim,¹ Hoon Ko² and Sung Bum Pan³

¹Department of Control and Instrumentation Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 61452, Republic of Korea

²IT Research Institute, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 61452, Republic of Korea

³Department of Electronics Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 61452, Republic of Korea

Correspondence should be addressed to Sung Bum Pan; sbpan@chosun.ac.kr

Received 14 December 2018; Accepted 17 January 2019; Published 3 February 2019

Guest Editor: Kwangjin Park

Copyright © 2019 Min-Gu Kim et al. 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.

IoT enabled smart car era is expected to begin in the near future as convergence between car and IT accelerates. Current smart cars can provide various information and services needed by the occupants via wearable devices or Vehicle to Everything (V2X) communication environment. In order to provide such services, a system to analyze wearable device information on the smart car platform needs to be designed. In this paper a real time user recognition method using 2D ECG (Electrocardiogram) images, a biometric signal that can be obtained from wearable devices, will be studied. ECG (Electrocardiogram) signal can be classified by fiducial point method using feature points detection or nonfiducial point method due to time change. In the proposed algorithm, a CNN based ensemble network was designed to improve performance by overcoming problems like overfitting which occur in a single network. Test results show that 2D ECG image based user recognition accuracy improved by 1%~1.7% for the fiducial point method and by 0.9%~2% for the nonfiducial point method. By showing 13% higher performance compared to the single network in which recognition rate reduction occurs because similar characteristics are shown between classes, capability for use in a smart vehicle platform based user recognition system that requires reliability was demonstrated by the proposed method.

1. Introduction

Recent active convergence between IT and industry has resulted in various cutting edges, IT being grafted in cars, and this has led to progress in development of smart cars with improved driver safety and convenience. Smart car utilizes technology that maximizes safety while pursuing driver convenience by applying IoT (Internet of Things) technology [1]. While the concept of car IoT has not yet been clearly fixed, it can be defined to be a mobile connectivity based computing environment on which various customized user service industry can be created in addition to promoting traffic safety and reducing congestion. Along with wearable devices, the smart car industry is part of the next generation growth industry and supports state of art technology like advanced driver-assistance systems (ADAS) and connectivity. In addition, as shown in Figure 1, expansion of IoT and wearable technology, for which improvement

is based on user information, has allowed application of related technology to smart cars. If user specific information gathered from the wearable device worn by the driver can be shared with the smart car, it will be possible to provide more convenience functions in the car. For example, in addition to executing commands via user voice recognition, there is the drowsiness detection service that alerts the driver according to fatigue and risk of driving while being drowsy by checking the biometric signals and vehicle information. In order to provide such services, a system to analyze wearable device information on the smart car platform needs to be designed [2].

In this paper, we will study a deep learning based user recognition method in order to apply biometric signal obtained from wearable devices to a smart car platform. Deep learning is a method in which optimum values are output from various deepened layers that were added to the multilayer perceptron used in existing machine learning.

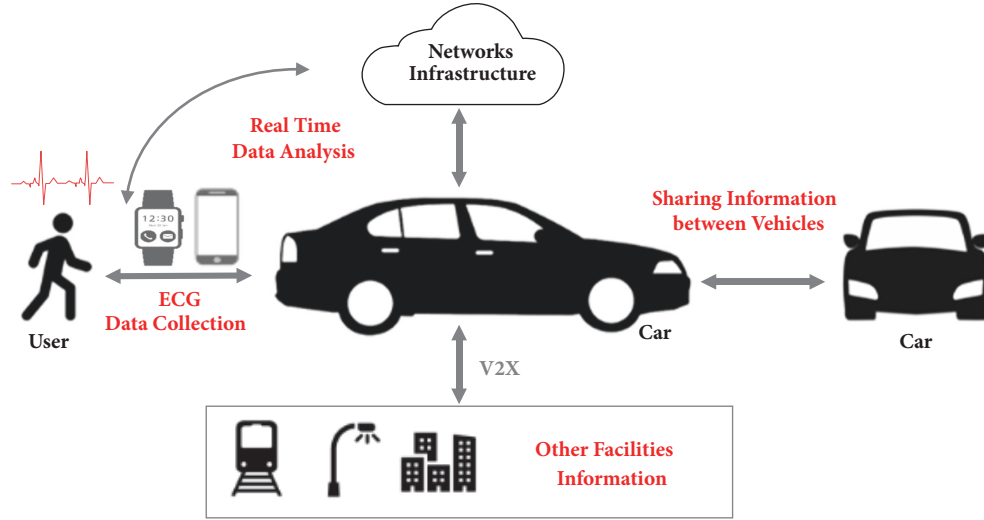


FIGURE 1: Structure of intelligent platform based smart car combined with IoT technology.

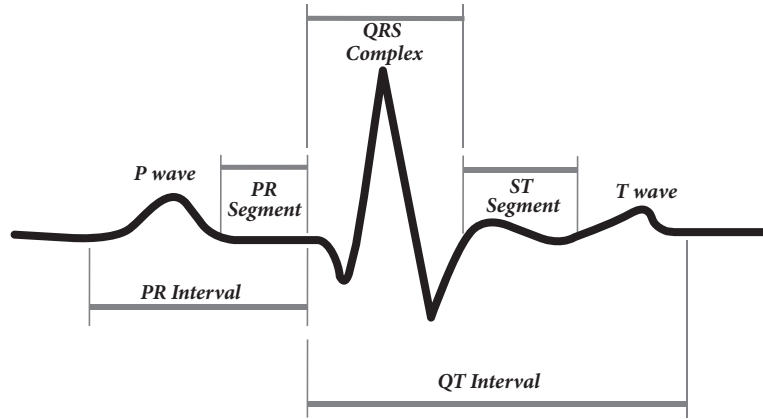


FIGURE 2: Example of ECG feature waveform.

Deep Neural Networks [3], Convolutional Neural Network [4], Recurrent Neural Network [5], Deep Belief Network [6], and Deep Q-networks[7] are some representative examples. In particular, the Convolutional Neural Network which has displayed outstanding performance in various areas including recognition, classification, and prediction is a neural network that has been designed to learn and classify various feature extraction filters automatically. While existing feature extraction method requires design by humans and only the classification is carried out through machine learning, feature extraction and classification are done automatically by the computer for the Convolutional Neural Network.

Existing user recognition methods that analyze physical features like the face, fingerprint, and iris were perceived to be safe as well as convenient. However, these methods use anatomical and physical form of information that is displayed externally and requires cooperation from the user and face to face type user intervention. Additionally, since such information can be forged or altered, problems can occur if applied to services that require a high level of security

like fintech, smart medicine, health-care and smart car service. Research in user recognition using biometric signals like ECG, EEG (Electroencephalogram), and EMG (Electromyogram) is being carried out to solve such problems. Important advantages of user recognition using biometric signals are as follows. (1) First, compared to methods that use anatomical and physical form of information that is displayed externally like face and finger print, this method uses signals that occur internally and are hard to forge, (2) they can be obtained from all living people, (3) they include information on clinical and psychological state of the user, and (4) finally, user rerecognition is easy to carry out since the waveform does not change by much over time [8]. Since ECG is a biometric signal that is nonreactive and hard to alter, it is being studied as a next generation user recognition technology. As shown in Figure 2, ECG is different for each individual according to factors like the location, size and structure of the heart, age, and gender. Normal ECG signal contains specific feature points and they can be used to extract features for user recognition. Therefore, an individual can be

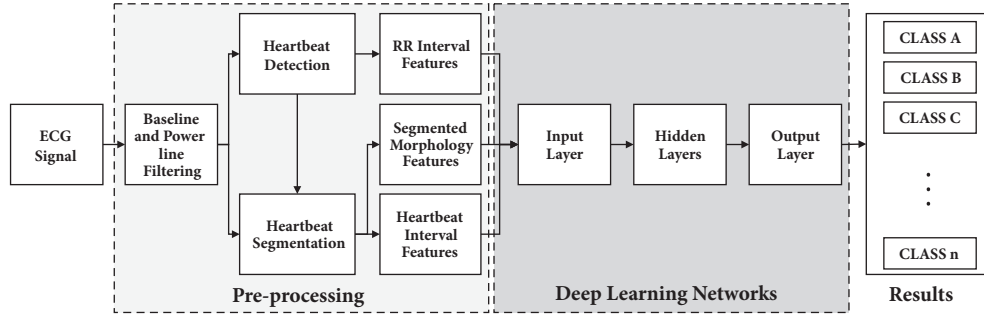


FIGURE 3: Flowchart of existing deep learning for ECG classification.

identified by using features unique to each individual and the user can be recognized by utilizing the ECG signal which can be measured no matter where the individual is located.

In this paper, existing research on ECG based user recognition will be analyzed in Section 2 and the proposed deep learning based ensemble network using ECG data will be explained in Section 3. Section 4 will contain analysis of performance results for the proposed method and the conclusion and suggestion for future research will be presented in Section 5.

2. Related Work

Israel proposed an ECG based recognition system using temporal features. After removing noise from the input ECG signal, P, QRS and T waveforms were detected and classification was carried out using Linear Discriminant Analysis after 15 features were extracted. Test on 29 subjects showed 100% subject recognition rate and 82% ECG beat recognition rate [9]. Biel proposed a classification method which compares test data consisting of 30 extracted fiducial point features with the trained group data and selected the best matching class using a Soft Independent Modeling of Class Analogy (SIMCA) classifier.

Test was carried out with a total of 20 subjects and 100% recognition rate was achieved. Even though this method showed a high recognition rate during testing, it is unsuitable for a real time recognition system since it extracts many features during the classification stage [10]. Wang proposed a method that uses a combination of time, amplitude, and R waveform features from the ECG signal. Fiducial point detection was carried out on preprocessed ECG signal in order to measure the time and amplitude distance, and Morphological Characteristics were extracted from main components using Linear Discriminant Analysis. When classification was carried out after two types of features were combined, 100% subject recognition rate and 98.9 recognition rate were achieved for 13 subjects. While a relatively high recognition result was obtained, it was from a small sample of subjects [11]. Shen did research in which overall recognition rate between past data and data obtained by repeated measurement of 23 identical subjects was compared. Test result from template matching that compared morphological difference of the ECG signal by extracting 17

fiducial features showed recognition rates of 98.5% and 87.7%. However, even though the matching result had decreased, the recognition system using the ECG signal was found to display high performance even after some time has passed if calibration is carried out [12]. Chan proposed a feature extraction framework using a distance measurement set that includes the wavelet transform distance. Data was obtained from 50 subjects through electrodes placed between fingers. By applying the wavelet transform distance method for the user recognition method, 89% recognition rate was achieved. Even though test was carried out on many subjects compared to previous tests, it showed a low recognition rate [13]. Chiu proposed an individual identification method using the wavelet and Euclidean classifier. After acquiring ECG signal from 45 subjects, features were extracted using wavelet transformation and the Euclidean distance was utilized [14]. While a relatively simple process was carried out, a low 90.5% recognition result was shown. Loong used Linear Predictive Coding (LPC) parameters for ECG signal classification [15].

As shown in Figure 3, for research that applied the ECG signal to deep learning, Rajpurkar developed a new network composed of 34 layers that forecast 12 variety of arrhythmia from a single lead ECG signal [16] and Chauhan developed a neural network structure that uses many LSTM layers repetitively in order to detect irregular ECG signals [17]. Rahha studied a neural network structure consisting of feature representation layer and Softmax regression layer. In order to remove accumulated noise from the ECG signal, the feature representation layer classified the ECG signal through the Softmax regression layer after training with autoencoders, a nonsupervised learning method [18]. The recent research on user recognition methods using deep learning in various technology fields such as recognition, classification, and prediction has shown excellent performance as shown in Table 1.

However, existing neural network composed of a single structure network has a performance limiting problem since one network cannot learn even data that is difficult to recognize. If learning is carried out for even data that is difficult to recognize, performance degradation will be the result due to overfitting that occurs during the training process [19]. In this paper, a deep learning based ensemble network that relearns good features that were output from many networks for data that is difficult to learn by a single

TABLE 1: Performance analysis of previous research on user recognition using ECG.

Works	Feature	Result	Number of Subject
Israel	LDA	ACC: 100%	27
Wang	PCA, LDA	ACC: 98%	13
Chan	WT	ACC: 89%	50
Chiu	WT	ACC: 90.5%	45
Louis	LBP, SVM	EER: 7.89%	1,020
Hejazi	WT	FMR: 3.5%	52

network is proposed. The algorithm proposed in this paper designs an ensemble network in order to solve the recognition rate reduction problem caused by exhibiting similar features between classes and the overfitting that occurs in existing single structure network. Relearning is carried out by blending good features that are output from each single network. In addition, feasibility of use in a wearable device based real time recognition system is verified by applying a 2D ECG image to the CNN model which has shown superior performance for image recognition and classification. ECG data used in testing was obtained from 18 test subjects in the MIT-BIH Normal Sinus Rhythm Database (NSRDB). 80% of the data was used as training data and the remaining 20% was used as test data. Test for 2D ECG image that was classified by the fiducial point method resulted in 97.3%, 97.9%, and 97.2% recognition rates for the case in which a single network was used for the 3 networks designed in the ensemble network while 98.9% recognition rate was obtained for the case in which the ensemble network was used, showing an increase in performance of 1.7%. For the case in which classification was carried out by the nonfiducial point method, single network displayed recognition rates of 97.2%, 96.6%, and 97.7% while the ensemble method resulted in recognition rate of 98.6%. Best results were achieved when the ensemble network was used for both the fiducial point method and the nonfiducial point method.

3. Proposed Algorithms for Ensemble Networks Based on ECG Data

Figure 4 shows the flowchart for the ensemble network based user recognition system using the ECG signal that is proposed in this paper. First, noise that occurs during ECG signal acquisition is removed through frequency filtering for DB formation. After detecting the R waveform peak using the Pan&Tompkins method, signal segmentation for each period including P, QRS, and T waveform is carried out. However, baseline fluctuation noise caused by respiration of person carrying out the measurement is not removed by the filter and median filter and the ECG image is acquired by estimating the partial baseline using first order regression analysis and projecting on the 2D space.

3.1. ECG Data Preprocessing. Preprocessing consisting of noise removal and calibration is carried out to remove noise and distortion before transforming 1D ECG signal into 2D image. ECG signal noise is removed via frequency

filtering, R waveform detection, and median filtering process as shown in Figure 5 [20]. A band pass filter was used for frequency filtering in order to remove power line noise, muscle noise, and electrode contact noise which occur during ECG measurement. R waveform peak was detected from the ECG signal that had passed through the band pass filter using the Pan&Tompkins algorithm. Using R waveform peak as a fiducial, noise was removed by applying the median filter on the remaining interval excluding the QRS Complex interval which includes the physical feature information unique to each individual.

However, a calibration process is needed since baseline fluctuation noise caused by user respiration is not removed by filter and median filter. For the technology used to remove the baseline fluctuation noise, partial baseline was estimated by using first order regression analysis.

$$\text{Partial baseline} = w \cdot x + \beta,$$

$$w = \frac{y_2 - y_1}{x_2 - x_1}, \quad \beta = y_1 - \frac{y_2 - y_1}{x_2 - x_1} \cdot x_1 \quad (1)$$

For the partial baseline, first order regression analysis is calculated by applying $M_{TP}(x_1, y_1)$, the median value between the ECG T waveform and P waveform, and $M_{PQ}(x_2, y_2)$, the median value between the ECG P waveform and Q waveform to (1). Finally, 1D ECG signal is transformed into 2D image through projection and linear equation for application to 2D-CNN. 1D ECG signal is projected using the amplitude value due to temporal fluctuation in (2).

$$PS_{ECG} = A_{ECG}(t) + |A_s| + \alpha \quad (2)$$

PS_{ECG} is the ECG pixel position and A_{ECG} is the ECG amplitude value at time t . 1D ECG signal is a discontinuous voltage value in the time domain and data loss will occur between pixels when projected on 2D space. Therefore, after minimizing data loss through linear equation, fiducial point based ECG signal was classified by projecting 1D ECG signal onto 2D image space using (3) as shown in Figure 6(a).

$$\begin{aligned} T \left(R_{peak} (N - 1) + \frac{d}{2} \right) &\leq T(N) \\ &\leq T \left(R_{peak} (N + 1) + \frac{d}{2} \right) \end{aligned} \quad (3)$$

Nonfiducial point based ECG classification is carried out with a fixed length sliding window as shown in Figure 6(b). For t

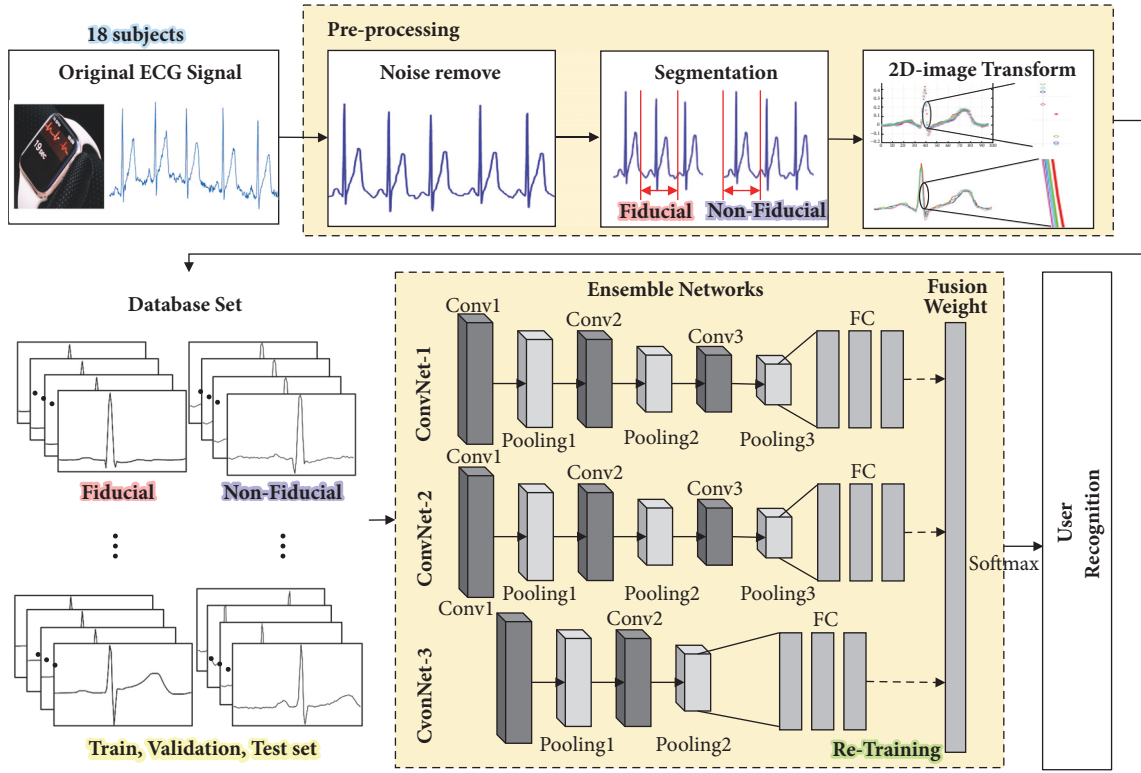


FIGURE 4: Flowchart of proposed user recognition system using ECG.

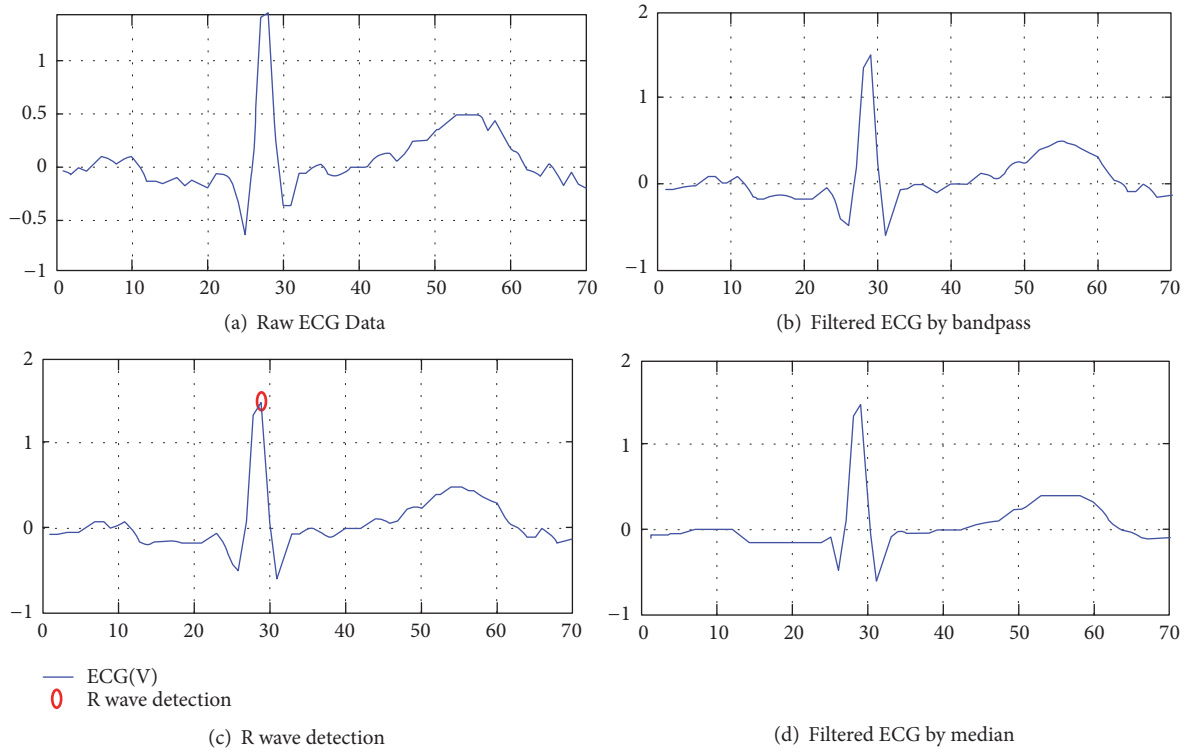


FIGURE 5: Noise removal process.

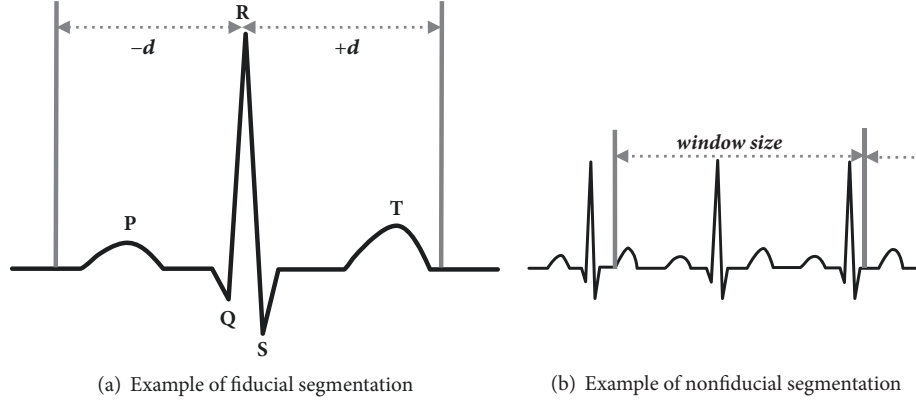


FIGURE 6: ECG signals segmentation methods.

test subjects, n signals to be classified for each test subject and d as the length of the signal to classify, data of size $n \times d$ can be obtained. Since the entire ECG for one person is longer than d , classification into many ECG signals is possible. The i th ECG signal is expressed as $x_{i,1:t} = \{x_{i,1}, \dots, x_{i,t}\}$, and the j th sample of $x_{i,1:t}$ is expressed as $x_{i,j} = [x_{i,j,1}, \dots, x_{i,j,d}]^T$. Since nonfiducial point based ECG uses all acquired signals, it was possible to construct a large amount of data compared to the fiducial point based ECG classification that only uses data between P, QRS, and T waveforms.

3.2. Ensemble Architecture Based on Deep Convolutional Neural Networks. Figure 7 shows the entire structure of the ensemble based CNN model proposed in this paper. CNN learns by extracting the features autonomously using a fixed size filter on information that is continuously input. Rather than using the input information as for feature extraction, only meaningful specific information is extracted and learned thereby reducing the volume of complex computation that occurs during the training process. Many features can be extracted by varying the filter size and features extracted using one or many filters make up one layer. In addition, data augmentation is applied in order to increase the limited data amount since recognition performance improves as the amount of data used in learning is increased. Since data loss and deformation are severe when there is left right reversal, scaling and color alteration are applied to 2D ECG and data amount can be increased by using up, down, left, and right translation on the image location with the original image as the fiducial as shown in Figure 8.

Recently introduced ResNet[21] and DenseNet[22] were proposed in order to solve the problem in which initial features of the image are lost at the final output stage as the depth of the CNN model is increased [23]. However, the ECG image consists of simple background and waveforms unlike regular image that contains complex patterns. Therefore, there is no need to design deep layer stages for the CNN model. On the contrary, performance can be degraded as the number of parameters is increased. Therefore, the proposed ensemble network consists of 3 single CNN models as the other layer number and parameter values. Good features extracted from

each network are used for training by blending into one data. Nonlinear activation functions like Rectified Linear Units (ReLU), Leakage Rectified Linear Units (LReLU), and Exponential Linear Units (ELU) are used in the CNN model to output kernel weight. Since ReLU that is most used in CNN converts negative numbers to 0, ELU which allows use of negative numbers was used.

$$\text{ReLU}(x) = \max(0, x) \quad (4)$$

$$\text{LReLU}(x) = \max(0, x) + \alpha \min(0, x) \quad (5)$$

$$\text{ELU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \gamma(\exp(x) - 1) & \text{if } x < 0 \end{cases} \quad (6)$$

An optimization function was used for the cost function, since it measures the learning state of the CNN model and displays the difference between the sample used for training and the expected output data Adam, a gradient descent based optimization function, was used to minimize the cost function. While ConvNet-1 and 2, the single networks used in the ensemble network, are made up of identical layer stages consisting of 3 Convolution Layers, 3 Max-Pooling Layers, and 3 Fully Connected Layers, differentiation was achieved by using Learning Rates of 0.001 and 0.01, respectively. ConvNet-3 consists of 2 Convolution Layers, 2 Max-Pooling Layers, and 3 Fully Connected Layers and 0.001 was used as the Learning Rate.

4. Experimental Results

The ECG database was obtained by taking 128 sampling measurements each from 18 subjects consisting of 5 males (ages 26~45) and 13 females (ages 20~50) in the MIT-BIH Normal Sinus Rhythm Database (NSRDB), an ECG DB. In this paper, 4,500 pages of training data, 2,700 pages of verification data and 1,800 pages of test data were used for the fiducial point based ECG image data and 90,000 pages of training data, 54,000 pages of verification data and 36,000 pages of test data were used for the nonfiducial point based ECG image data. For performance analysis of each class, Precision, Recall, F1-score, and Accuracy, the performance evaluation standards

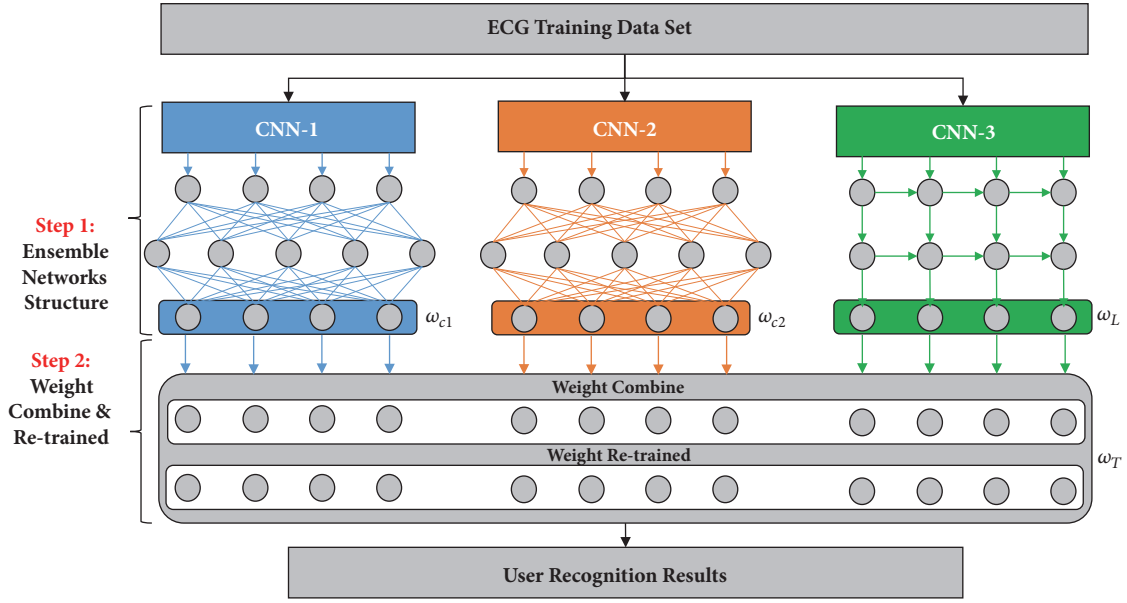


FIGURE 7: Structure of proposed ensemble networks.

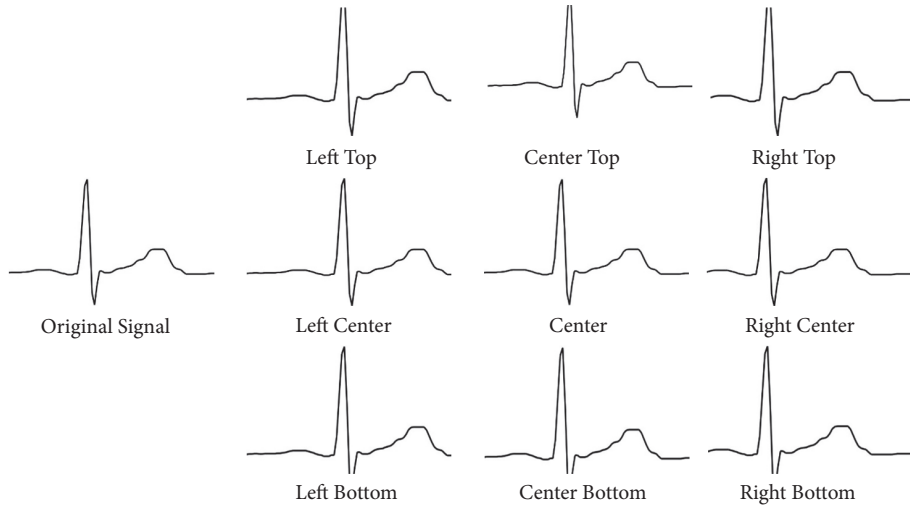


FIGURE 8: Data augmentation for ECG data increasing.

used in pattern recognition, were used. Precision, Recall, F1-score, and Accuracy for each class were calculated by applying TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) in (7), (8), (9), and (10).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

$$\text{Accuracy} = \frac{\text{Total correctly classified signals}}{\text{Total number of signals}} \times 100 \quad (10)$$

Figures 9 and 10 show the result displayed using the confusion matrix of the user recognition performance for fiducial point and nonfiducial point based 2D ECG image applied to the ensemble network. It is the result obtained with the estimated value by entering the fiducial point based ECG test data for 1,800 subjects and nonfiducial point based ECG test data for 36,000 subjects in the trained neural network model.

The columns display the Ground Truth and the rows display the number of ECG data for the proposed method.

True label	A	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	6	0	91	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	F	0	2	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	99	0	0	0	1	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
	L	0	0	0	1	0	0	0	0	2	0	97	0	0	0	0	0	0	0	0
	M	0	0	0	1	0	0	1	0	0	0	0	98	0	0	0	0	0	0	0
	N	0	5	0	3	0	0	0	0	0	0	0	0	90	0	0	0	0	2	0
	O	0	21	0	0	0	0	0	0	0	0	0	0	0	79	0	0	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		

Predicted label

(a) User recognition using ConvNet-1

True label	A	94	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	C	4	0	88	0	0	0	3	3	0	0	0	0	0	1	0	1	0	0	0
	D	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	2	0	0
	F	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
	N	0	1	0	2	0	0	0	0	0	0	0	0	0	95	0	0	0	2	0
	O	0	0	0	0	0	1	0	0	0	0	0	0	0	89	0	10	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		

Predicted label

(b) User recognition using ConvNet-2

True label	A	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	4	0	88	0	0	0	1	7	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	99	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	F	0	0	0	0	0	99	0	0	0	0	1	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	1	0	0	0	0	99	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	1	0	0	0	98	1	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
	N	0	8	0	3	0	0	0	0	0	0	0	0	86	2	0	0	1	0	0
	O	0	3	0	0	0	0	0	0	0	0	0	0	15	81	0	1	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		

Predicted label

(c) User recognition using ConvNet-3

True label	A	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C	5	0	89	0	0	0	1	5	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	E	0	0	0	0	99	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	F	0	1	0	0	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0
	G	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
	I	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
	J	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
	K	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
	L	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
	N	0	0	0	3	0	0	0	0	1	0	0	0	0	93	2	0	0	1	0
	O	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0
	Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		

Predicted label

(d) User recognition using ensemble networks

FIGURE 9: User recognition results based confusion matrix using 2D fiducial ECG images.

The following example will help in understanding the confusion matrix: the 15th row in Figure 9(b) which is O-class is classified into 89 ECG data (True Positive) that were recognized accurately as O-class and 11 ECG data (False Negative) that were falsely recognized as F-class and Q-class from a total of 100 data. In addition, 1,708 non-O-class ECG data (True Negative) that were accurately recognized as classes other than O-class and 2 B-class and C-class ECG data (False Positive) that were falsely recognized as O-class were confirmed.

Using such analysis method, Table 2 shows the result on user recognition performance using the fiducial point based 2D ECG image. User recognition accuracy for the ensemble network proposed in this paper displayed 1%~1.7% improved performance compared to the single network based user recognition performance. Table 3 shows the result on user recognition performance using nonfiducial point based 2D ECG image. As in the fiducial point based user recognition performance, user recognition accuracy for the ensemble network displayed 0.9%~2% improved

performance compared to the single network based user recognition performance.

In particular, since N-class and O-class ECG data displayed similar waveforms with the lowest amount of change in P, QRS, and T waveform distance values when compared to other classes, a low recognition rate was shown when a single network is used. However, the highest recognition rate of 96.5% was shown when applied to the ensemble network model as shown in Figure 11. When compared to the single network, difference in recognition performance of 4.5%~13% was displayed.

5. Conclusions

In this paper, an ECG data based ensemble network method for user recognition in smart car platform environment was proposed. Recognition performance was verified by applying 2D ECG image to the CNN model which has displayed excellent performance in image recognition and classification fields. Training data was transformed into 2D

A	1981	0	9	1	0	1	2	0	0	0	0	0	3	1	1	0	0	1
B	0	1930	0	18	0	9	0	0	26	8	4	0	0	1	4	0	0	0
C	4	1	1983	1	0	1	1	6	0	0	1	0	1	0	1	0	0	0
D	0	2	0	1972	0	3	0	0	12	2	0	4	2	3	0	0	0	0
E	0	0	0	11	1965	1	0	0	10	0	0	0	7	4	2	0	0	0
F	0	4	0	1	0	1980	5	1	3	5	0	0	1	0	0	0	0	0
G	1	12	10	0	0	16	1942	0	1	0	0	0	5	5	7	0	1	0
H	0	0	2	0	0	24	1	1973	0	0	0	0	0	0	0	0	0	0
I	0	2	0	6	2	0	1	0	1975	4	2	1	0	4	2	0	0	1
J	0	9	0	8	1	3	0	0	1	1957	7	0	6	1	3	0	0	4
K	0	3	1	5	0	0	1	0	2	1954	0	13	14	2	0	0	0	3
L	0	6	2	1	0	25	0	1	2	2	0	1941	2	2	16	0	0	0
M	1	8	0	4	0	6	0	0	5	1	1	1959	1	12	0	2	0	0
N	0	5	0	0	0	1	1	0	0	0	0	2	1979	10	0	1	1	1
O	0	4	31	9	0	5	19	17	0	1	0	11	161	8	1716	0	18	0
P	5	10	0	35	8	2	18	0	5	0	0	0	8	56	1853	0	0	0
Q	2	4	5	1	0	2	14	0	7	1	0	0	15	0	2	1	1946	0
R	0	4	3	0	0	0	1	0	1	1	1	0	1	4	2	0	0	1982
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	

(a) User recognition using ConvNet-1

A	1963	0	25	1	0	0	1	6	2	1	0	0	0	0	0	0	1	0
B	0	1763	0	20	0	8	0	0	175	4	8	0	1	2	13	0	0	3
C	0	0	1969	0	0	1	0	26	1	0	0	0	2	0	1	0	0	0
D	0	0	0	1872	0	1	0	0	121	2	0	2	0	2	0	0	0	0
E	0	0	0	0	1993	0	0	0	1	0	0	0	0	0	0	6	0	0
F	0	0	0	0	0	1961	1	1	35	1	0	0	1	0	0	0	0	0
G	0	2	21	0	2	13	1921	2	16	0	0	0	14	0	5	4	0	0
H	0	0	1	0	0	17	0	1981	0	0	0	0	1	0	0	0	0	0
I	0	0	1	0	1	0	0	0	1995	0	0	0	2	0	1	0	0	0
J	0	2	2	2	2	1	0	0	9	1970	7	0	2	0	2	0	0	1
K	0	1	0	3	0	0	0	0	0	2	1974	0	12	2	1	0	0	5
L	0	0	8	0	1	7	0	0	6	1	4	1947	11	1	13	0	0	1
M	1	0	0	0	1	3	0	0	14	4	3	0	1963	0	9	1	1	0
N	0	1	0	1	3	1	0	0	1	0	0	0	4	1845	142	0	2	0
O	0	0	19	0	4	1	12	0	7	0	2	4	128	1	1780	0	42	0
P	0	2	3	0	10	2	3	0	10	0	0	0	2	0	4	1964	0	0
Q	0	0	5	0	1	1	11	1	16	11	0	0	18	0	0	1	1935	0
R	0	0	6	1	0	0	0	0	1	0	0	0	1	0	0	0	1991	0
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	

(b) User recognition using ConvNet-2

A	1972	0	6	1	0	0	2	15	2	1	1	0	0	0	0	0	0	0
B	0	1862	0	30	0	1	0	1	91	8	4	1	0	0	1	1	0	0
C	1	0	1960	0	0	0	0	38	0	0	1	0	0	0	0	0	0	0
D	0	1	0	1981	0	1	0	0	13	0	0	4	0	0	0	0	0	0
E	0	0	0	7	1970	3	0	0	5	0	0	0	0	0	1	14	0	0
F	0	8	0	0	0	1935	3	1	47	2	0	0	4	0	0	0	0	0
G	0	3	1	0	0	6	1970	5	5	0	1	0	5	0	3	1	0	0
H	0	0	1	0	0	5	1	1993	0	0	0	0	0	0	0	0	0	0
I	0	0	0	3	0	0	0	0	1993	2	0	0	0	0	1	0	0	1
J	0	3	0	16	0	0	0	6	1967	2	0	3	0	0	0	1	2	2
K	0	1	0	4	0	0	3	0	1	2	1973	0	8	4	2	0	0	2
L	0	1	1	1	0	1	0	3	2	2	0	1980	1	1	6	0	0	1
M	1	2	0	0	0	4	0	1	2	2	0	1	1967	0	11	2	7	0
N	0	1	0	1	0	2	1	1	0	0	0	1	5	1972	11	0	5	0
O	0	1	5	1	0	1	10	6	1	0	0	8	109	2	1804	0	52	0
P	0	6	0	1	3	6	11	0	15	0	0	4	0	5	1949	0	0	0
Q	0	0	8	1	0	0	17	0	9	7	0	0	1	0	0	4	1953	0
R	0	1	8	2	0	0	0	1	1	1	0	0	6	0	0	0	4	1976
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	

(c) User recognition using ConvNet-3

A	1984	0	6	1	0	0	2	5	0	1	0	0	0	0	0	1	0	0
B	0	1916	0	5	0	3	5	0	57	9	3	0	0	1	0	0	0	1
C	0	1	1973	0	0	0	0	25	0	0	1	0	0	0	0	0	0	0
D	0	1	0	1974	0	1	0	0	3	17	1	0	1	1	1	0	0	0
E	0	0	0	0	1994	0	0	0	1	0	0	0	0	0	1	4	0	0
F	0	1	0	0	0	1986	4	0	8	1	0	0	0	0	0	0	0	0
G	0	1	4	0	0	7	1975	1	8	0	1	0	0	2	1	0	0	0
H	0	1	0	0	0	4	0	1996	0	0	0	0	0	0	0	0	0	0
I	0	1	0	0	0	0	0	0	1994	2	1	0	0	0	1	0	0	1
J	0	2	0	1	0	0	0	0	2	1993	2	0	0	0	0	0	0	0
K	0	1	0	2	0	0	0	0	2	2	1985	0	2	5	0	0	0	1
L	1	1	2	0	0	3	0	1	3	1	1	1981	1	2	2	0	0	1
M	1	4	0	0	0	3	3	0	2	5	1	0	1970	0	9	1	1	0
N	0	1	0	0	0	1	1	0	0	0	0	0	2	1992	3	0	0	0
O	0	4	16	0	0	0	12	1	1	3	8	6	70	8	1853	0	18	0
P	0	4	0	0	1	0	2	0	4	2	0	0	0	0	4	1982	1	0
Q	0	0	9	0	0	0	15	2	8	4	0	0	5	0	1	8	1948	0
R	0	0	3	1	0	0	0	0	1	0	0	0	2	2	0	0	0	1991
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	

(d) User recognition using ensemble networks

FIGURE 10: User recognition results based confusion matrix using 2D nonfiducial ECG images.

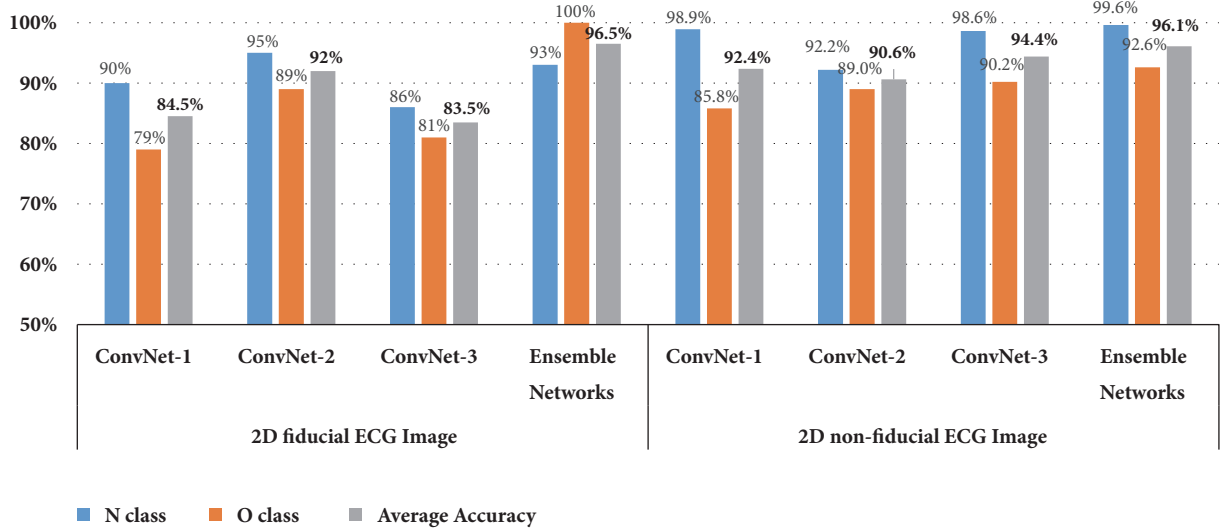


FIGURE 11: User recognition accuracy rate in similar feature class type.

image by proceeding through the process of noise removal and signal segmentation for each period after classifying them with the fiducial point based method using feature point detection and nonfiducial point method due to time change. In addition, an ensemble network was designed to

retrain with good feature data extracted from each single network in order to process data that is difficult to train in a single network and applied to the user recognition system. When the performance of the method proposed in this paper was analyzed, accuracy of the ensemble network based user

TABLE 2: Performance analysis of user recognition using 2D fiducial ECG images.

	Precision	Recall	F1-Score	Accuracy
ConvNet-1	0.98%	0.97%	0.97%	97.3%
ConvNet-2	0.98%	0.98%	0.98%	97.9%
ConvNet-3	0.97%	0.97%	0.97%	97.2%
Ensemble Networks	0.99%	0.99%	0.99%	98.9%

TABLE 3: Performance analysis of user recognition using 2D nonfiducial ECG images.

	Precision	Recall	F1-Score	Accuracy
ConvNet-1	0.97%	0.97%	0.97%	97.2%
ConvNet-2	0.97%	0.97%	0.97%	96.6%
ConvNet-3	0.98%	0.98%	0.98%	97.7%
Ensemble Networks	0.99%	0.99%	0.99%	98.6%

recognition using 2D fiducial point ECG image displayed improvement of 1%~1.7% compared to the single network based user recognition performance. User recognition result for 2D nonfiducial point ECG image using the ensemble network based user recognition method displayed 0.9%~2% improved performance compared to the single network based user recognition performance. In addition, by solving the problem inherent in single networks that display low recognition rate due to the inability to recognize similar waveforms with the proposed ensemble network, potential for application in a smart car platform that requires high reliability was verified. In the future, we plan to conduct research on user recognition for real life applications by acquiring multiple proprietary DBs based on state changes like the ECG measurement location and user workout, sleep and after drinking. In addition, we will also compare and analyze the performance of 1 dimensional ECG signals applied to the same network without converting them into 2D images. And we plan to improve the recognition performance by designing and combining various networks based on user change.

Data Availability

The Normal Sinus Rhythm DB data used to support the findings of this study have been deposited in the MIT-BIH repository PhysioNet.

Conflicts of Interest

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

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. 2018R1A2B6001984) and supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1A6A1A03015496).

References

- [1] J.-F. Bonnefon, A. Shariff, and I. Rahwan, "The social dilemma of autonomous vehicles," *Science*, vol. 352, no. 6293, pp. 1573–1576, 2016.
- [2] J. Fleetwood, "Public health, ethics, and autonomous vehicles," *American Journal of Public Health*, vol. 107, no. 4, pp. 532–537, 2017.
- [3] C.-H. Chang, "Deep and shallow architecture of multilayer neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2477–2486, 2015.
- [4] P. Le Callet, C. Viard-Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 17, no. 5, pp. 1316–1327, 2006.
- [5] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3639–3655, 2017.
- [6] W. Diao, X. Sun, X. Zheng, F. Dou, H. Wang, and K. Fu, "Efficient saliency-based object detection in remote sensing images using deep belief networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 137–141, 2016.
- [7] H. Sasaki, T. Horiuchi, and S. Kato, "A study on vision-based mobile robot learning by deep Q-network," in *Proceedings of the 56th Annual Conference of the Society of Instrument and Control Engineers of Japan, SICE 2017*, pp. 799–804, Japan, September 2017.
- [8] M. R. Ogiela and L. Ogiela, "On using cognitive models in cryptography," in *Proceedings of the 30th IEEE International Conference on Advanced Information Networking and Applications, AINA '16*, pp. 1055–1058, Crans-Montana, Switzerland, March 2016.
- [9] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133–142, 2005.
- [10] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.
- [11] Y. Wang, F. Agraftioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 148658, pp. 148–658, 2007.

- [12] T.-W. Shen, S.-C. Chang, C.-H. Wang, and T.-C. Fang, "A longitudinal and cross-section investigation on peritoneal dialysis patients: Does the cardiovascular conditions affect on ECG biometrics?" in *Proceedings of the Computing in Cardiology 2010, CinC 2010*, pp. 497–500, Belfast, Ireland, September 2010.
- [13] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 2, pp. 248–253, 2008.
- [14] C. C. Chiu, C. M. Chuang, and C. Y. Hsu, "A novel personal identity verification approach using a discrete wavelet transform of the ECG signal," in *Proceedings of the 2008 International Conference on Multimedia and Ubiquitous Engineering (mue 2008)*, pp. 201–206, Busan, Korea, April 2008.
- [15] J. L. C. Loong, K. S. Subari, R. Besar, and M. K. Abdullah, "A new approach to ECG biometric systems: a comparative study between LPC and WPD systems," *World Academy of Science, Engineering and Technology*, vol. 68, no. 20, pp. 759–764, 2010.
- [16] P. Rajpurkar, A. Y. Hannun, M. Haghpandhi, C. Bourn, and A. Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," 2017.
- [17] S. Chauhan and L. Vig, "Anomaly detection in ECG time signals via deep long short-term memory networks," in *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Paris, France, October 2015.
- [18] M. M. A. Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.
- [19] D. M. Hawkins, "The problem of over-fitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 1, pp. 1–12, 2004.
- [20] G. Choi, J. Jung, H. Moon, Y. Kim, and S. Pan, "User authentication system based on baseline-corrected ECG for biometrics," *Intelligent Automation and Soft Computing*, 2018.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pp. 770–778, July 2016.
- [22] G. Huang, Z. Liu, L. V. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700–4708, July 2017.
- [23] T. J. Jun, H. M. Nguyen, D. Y. Kang, D. H. Kim, D. Y. Kim, and Y. H. Kim, "ECG arrhythmia classification using a 2-D convolutional neural network," in *Proceedings of the Computer Vision and Pattern Recognition*, 2018.

