Untact Abnormal Heartbeat Wave Detection Using Non-Contact Sensor through Transfer Learning

JIN-SOO KIM AND KANGYOOK LEE
Department of Computer Engineering, Gachon University, Seongnam-Si 13120, South Korea
Corresponding author: KangYoon Lee (keylee@gachon.ac.kr)

ABSTRACT
This paper presents an important advancement in heart activity monitoring, focusing on non-contact sensor data, which tend to be noisy due to interference, and the limitations of non-contact (untact) technology. A preprocessing filter and optimal classification model are proposed to improve the accuracy and reliability of heart rate data measured by a non-contact Doppler radar sensor, and the results are compared to those of a contact heart rate sensor (Holter monitor). The MIT-BIH Arrhythmia Database of PhysioNet are used for learning, and the results from the non-contact sensor and Holter monitor are compared for verification. To train the abnormal heartbeat waveform classification model, (1) an optimal heart rate data separation window size is selected through iterative model comparison and used for data separation, and (2) meaningful indicators of heart rate variability are selected; the data are transformed and applied as model characteristics. The non-contact sensor data are then applied to three filter algorithms, and the accuracy is assessed by comparison with the contact sensor data using the trained abnormal heartbeat waveform classification model. Learning is performed using 12 classification models, and the accuracies of the models are compared. This study demonstrates an effective new method of transfer learning for contact data abnormality detection.

INDEX TERMS
Abnormal detection, classification model, heartbeat waveform, heart rate, non-contact sensor, preprocessing filter algorithm.

I. INTRODUCTION
The COVID-19 pandemic has led to a rise in demand for non-contact (untact) technology, including non-contact sensors. However, it is difficult to achieve the same data accuracy as contact sensors through non-contact sensors. The combination of Internet of Things (IoT)-based sensors and deep learning can help drive the use of untact technology for ambient intelligence (AmI), especially in the field of healthcare by means of unobtrusive biomedical monitoring of heart rate, respiration, etc. Heart rate abnormalities detected through heart rate data analysis can effectively reveal the condition of a subject before any complications occur [1]. Many studies have been conducted till date using various heart rate sensors, such as electrocardiography (ECG) and photoplethysmography (PPG) sensors, which are the most typical contact sensors. Several types of non-contact heart rate sensors, such as heat, light, or acceleration-based sensors, as well as Doppler radar sensors, have been used in many studies. However, the reliability of data measured using non-contact sensors is lower than that of data from contact sensors. Contact sensors have measurement units and methods different from those of non-contact sensors. In particular, the shape of the waveform measured by a non-contact sensor is highly irregular and sensitive to movement and environmental factors. Contact sensors can also show abnormal values based on movement and posture, rather than providing reliable data at all times. However, the reliability of beats per minute (bpm) data has been found to be insensitive to surrounding disturbances when compared with the raw bpm waveform data. The fact...
that contact and non-contact sensors both measure the heart rate in the same units of bpm enables comparison of the two types of sensors.

The contributions of this study are as follows. First, by using non-contact sensors, this study supports the accuracy of contact sensors by transfer learning model. Second, the model of the non-contact sensor is tested in a real-life situation. Unlike existing studies which have conducted learning and experiments with open heart rate data sets, in this study, an experiment is conducted and the data collected from the experiment is used in the learning model using an open dataset and the results are shown. An abnormal heartbeat waveform classification model is learned through an open database, and bpm data measured by contact and non-contact sensors are applied to the model. By applying this model to non-contact sensor, we confirm the possibility of untact processing. To improve the performance of the non-contact sensor, we propose and experiment with filter algorithms and optimal window size, and show meaningful results. The performance of a non-contact sensor is compared with that of a contact sensor, and a filtering algorithm is developed to improve the reliability of the non-contact sensor data. There have been many studies using the characteristics of heart rate variability (HRV). However, previous studies have used the redundant features of HRV for learning [2]. There is a limitation that only learning and testing through open datasets were conducted. This study has the distinction between increasing learning efficiency by reducing overlapping indicators and conducting experiments in untact environments through non-contact sensors. The results of this study will help promote the use of non-contact sensors in various applications. Non-contact sensors provide an excellent means of continuous biomedical data monitoring for doctors and health care professionals by enabling non-contact measurements during daily activities. Moreover, they will aid in the faster implementation of AmI in the field of healthcare [3].

II. RELATED RESEARCH

A. CONTACT SENSORS

Two typical types of contact sensors are ECG and PPG sensors [4]. An ECG shows a graphical representation of the electrical activity of the heart during the heartbeat cycle. Therefore, ECGs produce reliable data and are commonly used in clinical setting to diagnose arrhythmia based on irregular heartbeats [5]. They are also useful in identifying indicators of cardiac diseases as well as myocardial disorders, atrial hypertrophy of the atrial ventricle, and dilatation of pulmonary circulation [6], [7]. PPG acquires data by measuring the intensity of light reflected by blood flow, by illuminating the skin using a light-emitting diode [8]. The optical blood-flow sensor data provide various information such as cardiovascular information, arterial blood pressure, stiffness, pulse transition time, pulse rate, heart output, arterial compliance, and peripheral resistance, as well as heart rate. The contact sensor data used in this study were collected using a Holter monitor, which is an ECG-based medical device, and obtained from the open ECG data of PhysioNet [9].

B. NON-CONTACT SENSORS

There are various types of non-contact sensors depending on the measurement method [10]. One of the most typical types is the Doppler radar sensor, in which microwaves are used to detect changes in the period of the signal that is reflected back to the target. In addition to Doppler radar sensors, there are sensors to measure heart rate and pulmonary function using heat, light, and acceleration [11].

The non-contact sensor used in this study was a Doppler radar microwave sensor. Using this sensor, data, such as heart rate (bpm), respiration, and motion data, could be collected at a maximum distance of 1.2 m. The sensor data were captured by and stored in a Raspberry Pi 3. The data, including time and sensor measurements, were stored in two files. The first file contained the raw data measured by the non-contact sensor, which were unique values measured every 10 ms. A waveform could be generated from these data. The second file included the heart rate, respiration, and motion data in bpm. The bpm measurements were collected, and the average bpm value was measured every 5 s using three unique values: heart rate, respiration, and motion.

C. HRV

HRV is the best known means of assessing the function of the autonomic nervous system [12]. HRV is a periodic change in heart rate and quantifies the time between instantaneous heartbeats according to the change in the heart rate. After the QRS waveform is detected in ECG data, measurement is performed based on the R-R interval. Although there are many reasons for using HRV, the main reason is that there is no risk to the subject due to its non-invasiveness and the simple measurement technique. HRV investigations began to develop when an autonomic nervous system evaluation report was published in 1997 in accordance with the standards established by an international expert committee composed of members of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology in 1996 [13]. Two methods can be used for HRV assessment: the time-domain and frequency-domain methods. The time-domain method involves measuring the change in heart rate during a normal period or predetermined time, and the frequency-domain method involves measuring the periodic vibration of heartbeat signals acquired by decomposing the heart rhythm signal into various frequencies and vibrations [14].

III. CLASSIFICATION MODELING PROCESS

The abnormal heartbeat waveform classification modeling process in this study is divided into five parts, namely, data acquisition, data preprocessing, data conversion, model learning, and model validation, as shown in Fig. 1. These processes involve identifying and comparing the applicability of non-contact sensor data to the abnormal heartbeat waveform.
A. DATA COLLECTION

Data collection, which is the first step in the abnormal heart rate classification modeling process, involves obtaining heart rate data using the contact and non-contact sensors. Heart rate data collection using the contact sensor includes open heart rate data acquisition and direct measurement through a Holter monitor. The open heart rate data were obtained from the MIT-BIH Arrhythmia Database of PhysioNet [14]. There are several types of ECG data in the MIT Arrhythmia Database, among which R-R interval (RRI) data is used in this study. Each RRI datum is labeled with the classification of the corresponding heart rate waveform. Heart rate data is simultaneously collected using the Holter sensor and non-contact sensor. A Holter electrode is attached to the body of the subject, and the heart rate is measured about 50 cm away from the front of the non-contact sensor. When the body of the subject moves during measurement, the non-contact sensor and Holter sensor data capture the variability. Therefore, minimal movement is required. The non-contact sensor heart rate data is stored on the Raspberry Pi 3, and the contact sensor data is stored on an SD card inside the Holter device. The Holter monitor records the heart rate every 3 s, whereas the heart rate measurement cycle of the non-contact sensor is 5 s. Direct measurements made using the contact sensor are used to verify the accuracy of the non-contact sensor data, and the non-contact sensor data are used as test sets for the models developed from the open data.

B. DATA PROCESSING

After data collection and measurement, preprocessing is performed. Because the open heart rate data are highly reliable, preprocessing is skipped for these data and is mainly applied to the direct heart rate measurements. The initial data preprocessing involves synchronizing the data from the contact and non-contact sensors because their measurement intervals are different, at 3 s and 5 s, respectively. Thereafter, the characteristics of the non-contact sensor are identified to filter the non-contact sensor heart rate data using three filters.

The first filter removes the data at the beginning and end of the measurement. The contact sensor is stabilized if the electrode is normally attached and measured within approximately 5 s, after which the heart rate (HR) is measured immediately. In contrast, the non-contact sensor waits for 40 s before data is measured as shown in (1). The Beginning-End filter in Fig. 2 shows that the initial and final data measured by the non-contact sensor are abnormal values.

\[ T_{\text{time}} = \{1, 2, 3, 4, \ldots, n\}, \]
\[ HR_T = 0 \left( T < 40 \text{ or } T > n - 40 \right) \quad (1) \]

The second filter removes the outliers by setting a threshold. Exceptionally high values are removed by setting an upper threshold. Each outlier datum is replaced with the average of the data values immediately preceding and following it, rather than simply removing the datum as shown in (2). As shown in Fig. 2, 127 bpm is set as the threshold. 127 bpm corresponds to the largest value when compared with the measurement data of the contact sensor in the same environment; the value is found to be accurate during the data collection process.

\[ \text{Threshold} = E, \]
\[ HR_T = \frac{1}{n} \sum_{i=0}^{n} HR_{\text{peak}} \left( HR_T > E \right) \quad (2) \]

The third filter removes outliers based on successive numerical values. The measured HR data are time-series data with a period of 5 s in the case of the non-contact sensor. Fig. 2 shows a sudden change in the continuous time-series data, which can be found using successive numerical filtering. This filter converts the data into average data for consecutively measured periods. First, the line segments of the peak and valley points of the heartbeat are connected within 5 seconds. It then converts when it exceeds the sum of the peak and valley point of the next 5 seconds and the previous 5 seconds as in (4). In the filtering process, the data passing through two or more filters are immediately converted, and those passing through one filter are converted following review. However, the data filtered out by the threshold in the threshold filter are removed immediately because normal data cannot be obtained.

\[ X_T = \frac{HR_{\text{peak}} HR_{\text{valley}}}{X_{T-1} + X_{T+1}} \quad (3) \]
\[ X_T > X_{T-1} + X_{T+1} \]
\[ HR_T = \frac{1}{n} \sum_{i=0}^{n} HR_{\text{peak}} \quad (4) \]

C. DATA CONVERSION

The open heart rate data and directly measured data are converted into a data type that can be used for learning. Here, the open heart rate data are RRI (interval between R
waveforms during QRS waveforms) data and the measurement data are heart rate (bpm) data. Since the HRV used for learning is measured using RRI data, the measured heart rate data are converted into RRI data.

After the data is converted into RRI data, the optimal window size is found by separating the RRI data, which are time-series data, according to time such that the learning accuracy is maximized. The model is repeatedly trained to find the optimal window size. The window size is determined by the characteristics of the heart rate. The average human heart rate is 60 to 100 beats per minute, which converts to 1.33 seconds to one heartbeat. The change in the RR interval of the QRS waveform of HRV shows the highest change in 30 second units as shown in the experimental results in Table 1, which shows the learning accuracy with window sizes of 10, 15, 20, 25, 30, 35, 40, 45, and 50 s after performing learning using support vector machine (SVM), decision tree, k-nearest neighbor (KNN), and ensemble methods. It is based on 10-fold cross validation, and the optimal parameters are handled empirically using grid search. Although there is no significant difference in accuracy between most of the window sizes, the SVM (kernel function = 'polynomial,' C = 1, gamma = 2.5) and ensemble learning (AdaBoostM2) methods yield more than 90% accuracy with a window size of 30 s. After identifying the optimum window size as 30 s, the direct contact and non-contact sensor measurements are separated around 30 s.

After the data are separated, they are converted according to the HRV criteria shown in Table 2. Among these criteria, MRR is the mean of the currently separated RRI data; MNN is the mean of NN values, where NN is the difference in the RRI data, RRI(2)–RRI(1); and rMSSD is the square root of the squared mean of the consecutive NN differences. Furthermore, the NN50 criterion, which is the number of cases with a difference between NN intervals greater than 50 ms, or geometric analysis using a transformation method based on the fast Fourier transform in the frequency domain are included.

Dividing the open heart rate data using a window size of 30 s yields 3717 data points, which are converted into 16 heart rate variance criteria and used as features for model training. Each annotation is used as a label (y). Directly measured contact and non-contact data are also obtained from each of the 50 data sets and used as a test set.

D. ABNORMAL HEARTBEAT WAVEFORM CLASSIFICATION MODEL TRAINING

The model training is performed using an open heart rate database [2]. The open heart rate data are divided into 3717 data points with an optimal window size of 30 s, and the training and test sets are separated in an 8:2 ratio. The features used to train the model are based on 16 HRV criteria, and the classification categories are based on 19 heart rate classification criteria annotated with open heart rate data, as shown in Table 2. These criteria can produce reliable data that are collected and annotated directly by PhysioNet specialists who specialize in the distribution of biometric open data. However, only five classification categories are used in this study: atrial premature beat (A), left bundle branch block beat (L), right bundle branch block beat (R), premature ventricular contraction (V), and paced beat (/). There are three reasons for using only five classification categories.

The first and most important reason is data relevance. Heart conditions can be categorized using a wide variety of criteria, including numerical and waveform shapes. However, this study is based on RRI data. Therefore, when classifying the heart rate, the RRI is selected since it is the classification category most closely related to the heart rate data.

The second reason is the frequency. The five heart rate classification categories selected are generally the top five of the 19 categories in terms of identifying heart rate.

The third reason is realistic classification accuracy. The heart has been studied for a long time and in various ways, producing hundreds or thousands of related variables. To enable realistic classification using the current model, only five categories are selected. Various algorithms have been

---

**TABLE 1. Accuracy according to window size.**

| Window Size | SVM | DT | KNN | Ensemble |
|-------------|-----|----|-----|----------|
| 10          | 87% | 85%| 74% | 88%      |
| 20          | 83% | 84%| 79% | 85%      |
| 30          | 90% | 90%| 83% | 92%      |
| 40          | 85% | 86%| 82% | 88%      |
| 50          | 84% | 86%| 84% | 86%      |

---
studied in order to find an algorithm suitable for a given purpose. In this study, it is supervised learning for the purpose of classifying abnormal heartbeat waveforms using HRV indicators, and anomaly detection for determining abnormality indicators. Therefore, representative algorithms SVM, KNN, DT, and ensemble that satisfy these two conditions are used [15]. The SVM algorithm is divided into linear, secondary, tertiary, and Gaussian kernel SVM approaches according to the type of kernel, and the KNN algorithm is divided into cosine, tertiary, and weighted KNN methods. Finally, for the ensemble algorithm, learning is conducted using boosting and bagging techniques.

E. VALIDATION OF ABNORMAL HEARTBEAT WAVEFORM CLASSIFICATION MODEL

The open heart rate data are separated using the optimal window size, and the RRI of the separated open heart rate data is converted according to the HRV criteria. Thereafter, various classification models are trained using several algorithms and the most accurate algorithm is selected. Then, the measured heart rate data from the contact sensor, the Holter monitor, and from the non-contact sensor are separated according to the window size using the same process that is employed for the open heart rate data, the data are converted into HRVs, which are applicable to the trained model. For verification, the heart rate data measured by the Holter monitor and the non-contact sensor in the same environment are applied to the trained model, and the accuracy of the non-contact sensor data is compared to that of the Holter monitor data.

IV. EXPERIMENTAL IMPLEMENTATION

The sequence of the abnormal heartbeat waveform classification model experiment is shown in Fig. 3. Firstly, the experimental environment is set up and the sensor specifications are set. Secondly, the data measurement methods are implemented, and the accuracy of the data trained through each classification algorithm is assessed. The confusion matrix is obtained for the model with the highest accuracy for each classification algorithm, and these matrices are compared to select the final model. Finally, the measured contact and non-contact sensor heart rate data and before and after pretreatment are compared.

A. SENSORS

The environment includes the specifications of the contact and non-contact sensors and the machine learning environment. The contact sensor is a Holter ER-2000, whose specifications are shown in Table 3. A bipolar one-channel method is utilized, and measurement is performed directly on the body. The measured data are stored on the SD Card of the Holter device in the ECG file type. The non-contact sensor is a Sharp microwave Doppler radar sensor. The maximum sensing distance of the non-contact sensor is 1.2 m. Eight element transmission and detachable flat antennas are attached. The sensor has dimensions of 30 mm $\times$ 46.5 mm $\times$ 5 mm, a supply voltage of 3.7 V, and a typical current consumption of 80 mA. Images and specifications of the contact and non-contact sensors are shown in Fig. 4 and Table 3. The non-contact sensor developed by Sharp is being researched for use in the aging social medical patient monitoring project in Japan (i.e., the building of a nursing system based on non-contact heart rate and respiration measurement technologies). For comparison, the reliability of the Sharp non-contact sensor is evaluated by comparing it with that of a pulse oximeter from Pacific-Medico and a non-contact sensor produced by company A. Sharp non-contact sensors show higher performance than other non-contact sensors, as presented in Table 4. In the test environment, with the oximeter attached to the subject, the micro-sensors from company A and Sharp are
installed side by side and measurements are performed simul-
taneously. During the measurements, the displayed values are
collected and compared for subjects of various age groups.
The mean and variability of the heart rate and respiration
for each age group are closer to the results obtained by the
oximeter than those of the sensor from company A. The
average heart rate difference is 3.1, 2.1, 0.9, and 1.6 for the
subjects in their 20s, 30s, early 40s, and late 40s, respectively.
As shown in each figure, the Sharp non-contact sensor yields
values more similar to those of the oximeter than the other
non-contact sensor does, and the performance improves with
increasing age.

B. MEASURING HEARTRATE USING NON CONTACT SENSOR
The microwave sensor uses radio waves in the microwave
band 24 GHz band, and integrates the directional transmit
and receive antennas to form a Doppler sensor. As shown
in Fig. 4, there are installation methods that increase the
sensitivity when installing this sensor and follow Snell’s law
which is the reflection law of radio waves (light). The rec-
commended distance from the antenna surface of this sensor
to the human body / chest is about 60 cm to 100 cm. In the
lumbar area, it ranges from a few cm to 10 cm, in the back,
it ranges from several cm to 30 cm, and adjustments according
to the environment may be necessary. It is difficult to detect
pulse waves from the side of the human body, arms, legs,
etc. Fig. 5 shows an example of the distance characteristic
of heartbeat-respiration when microwave radiation is applied
to the chest from this sensor to a person in a sitting state
(after calibration is performed). This distance characteristic
has individual differences because the size of the chest and
the magnitude of the beat differ depending on the person.

C. TRAINING AND TESTING OF ABNORMAL HEART RATE
WAVEFORM CLASSIFICATION MODELS
The sequence of model testing and evaluation is as follows.
Firstly, the preprocessed open data are separated into training
and test sets, and various classification model algorithms
are applied for learning. In total, four classification model
algorithms are used: decision tree, SVM, ensemble, and KNN
algorithms [16]–[18]. By parameter learning and the various
methods involving these algorithms, an optimal classification
model is designed for each algorithm and the classification
model with the highest accuracy is selected. In the selected
model, the heart rate data measured simultaneously by the
Holter monitor and non-contact sensor are applied to the
trained classification model for accuracy comparison. The
accuracies of all the models used in the classification algo-

V. RESULTS AND DISCUSSION
A. DECISION TREE MODEL LEARNING TEST
In the decision tree algorithm, the Gini diversity index is set
as the partitioning criterion based on a maximum of 100 parti-
tions to perform the learning. Three kinds of tree structures—
dense, middle, and sparse trees—are studied in the learning,
and the dense tree structure produces the highest accuracy
at 84% (SplitCriterion = “gdi,” MaxNumSplits = “100,”
Surrogate = “off”). The accuracy of the middle tree structure
is slightly less, at 82%, whereas the coarse tree structure
produces the lowest accuracy of 67%. Based on the decision
tree confusion matrix in Fig. 6, the lowest accuracy among
the five anomaly classifications is found to be for category
A, with an accuracy of 61%.

B. SVM MODEL LEARNING TEST
The SVM classification model learning is also conducted in
three forms. One of the characteristics of the SVM approach
is that it can be used not only for linear models but also for
nonlinear data such as secondary and tertiary models. In this
study, linear, secondary, and tertiary SVM classification mod-
TABLE 3. Specification comparison of contact sensor and non-contact sensor.

| Contact Sensor | Non-Contact Sensor | Contact Sensor |
|----------------|--------------------|----------------|
| Induction method | Dipole 1 Channel method | 24 GHz u-wave Doppler method |
| Detection distance | Body contact | Max 1.2 m |
| Product size | 95 mm × 64 mm × 16 mm | 30 mm × 46.5 mm × 5 mm |
| Power | 3.7 V | 3.7 V |

TABLE 4. Comparison of non-contact sensor performance.

| Subject | Age | Product | Heart rate (average) | Respiration (average) | Heart rate (variability) | Respiration (variability) |
|---------|-----|---------|----------------------|-----------------------|--------------------------|--------------------------|
| A       | 28  | Oximeter Company A Sharp | 75.5 | 79.4 | 17.4 | 3.84 | 3.72 |
| B       | 39  | Oximeter Company A Sharp | 67.3 | 70.6 | 20.4 | 2.36 |
| C       | 40  | Oximeter Company A Sharp | 73.8 | 79.1 | 12.5 | 1.69 | 4.86 |
| D       | 49  | Oximeter Company A Sharp | 70.8 | 80.0 | 11.0 | 5.93 | 8.46 |

TABLE 5. Accuracies of all models used in the classification algorithm.

| Model       | Option | Accuracy |
|-------------|--------|----------|
| SVM         | Liner  | 82.7     |
|             | RBF    | 91.6     |
|             | Quadratic | 89.6    |
|             | Gaussian | 88.6    |
| Decision Tree | Fine  | 84.4     |
|              | Medium | 82.8     |
|              | Coarse | 64.6     |
| KNN         | Cosine | 86.3     |
|             | Coarse | 78.8     |
|             | Weighted | 80.5    |

els are trained. The secondary SVM model shows the highest accuracy of 91% (kernel function = “polynomial,” polynomialOrder = “2,” C = “1,” gamma = “2.7”), followed by the tertiary SVM at 89% (kernel function = “polynomial”) and the linear SVM (kernel function = “linear”) at 82%. The lowest is Gaussian SVM (kernel function = “rbf”). The confusion matrix of the secondary SVM model in Fig. 6 confirms that the error rate of category A is the highest, as in the decision tree case.

C. KNN MODEL LEARNING TEST

The KNN algorithm is also trained using three methods: the cosine, tertiary, and weighted KNN methods. The most accurate is the weighted KNN method, which produces an accuracy of 86% (Distance = “euclidean,” NumNeighbors = “10,” DistanceWeight = “SquaredInverse”), followed by the cosine KNN model at 80%, and the tertiary KNN model at 78%. In the confusion matrix of the most accurate tertiary KNN method in Fig. 6, the accuracies of A, /, and L are low, whereas those of the remaining classification categories, R and V, are high.

D. ENSEMBLE MODEL LEARNING TEST

The ensemble technique is classified into two categories—boosting and bagging tree approaches—for model learning. The learning rate is 0.1, and 300 iterations are performed. The accuracy of the ensemble boosting tree model after training is 91% (Method = “AdaBoostM2,” NumLearningCycles = “30,” LearnRate = “0.1”), and that of the bagging tree model is 90%. The ensemble technique is highly accurate as it uses a combination of several algorithms. In Fig. 6, all five classifications exhibit high accuracy in the confusion matrix corresponding to the boosting tree method.

E. DATA MEASUREMENT MODEL APPLICATION TEST

The above four algorithms are used to verify the accuracies of the 12 models generated. The confusion matrix, as shown in Fig. 6, of the model with the highest accuracy for each algorithm is selected as the final model, and accordingly, ensemble boosting model is selected. The preprocessing window size for the measured heart rate data is set to 30 s, and the HRV is converted using 16 criteria. Fifty heart rate data
F. RESULTS OF EXPERIMENT

In this experiment, heart rate abnormality waveform classification models are trained and applied to heart rate data from non-contact and contact sensors. Heart rate data for abnormal heart wave classification are used for learning, using HRV as the feature and 16 types of HRV indices. In the preprocessing before data conversion, training is performed by separating the time-series heart rate data according to the optimal window size for learning. Repeated experiments are conducted to determine the optimal window size, and the highest accuracy is obtained when the window size is 30 s. After machine learning is performed using heart rate data that has been converted into HRVs, the SVM and ensemble techniques produce the highest accuracy. After analyzing
the confusion matrices of these two classification models, the ensemble technique is selected because its accuracy is found to be well-balanced as shown in Table 6. The selected model is applied to the heart rate data measured using the non-contact sensor, yielding an accuracy of about 50% compared to the Holter sensor’s performance. After preprocessing using three filters based on the characteristics of the non-contact sensor, the accuracy increases to 75% with respect to the Holter sensor’ performance.

VI. CONCLUSION

The COVID-19 pandemic brings us to the untact environment sooner. An experimental method involving a non-contact sensor and a Raspberry Pi was developed. The method uses time-series HRV information, including RRI data, to provide a fast and resource-efficient approach that could be used on mobile and IoT devices. The open heart rate data obtained by the contact sensor were converted into HRV RRI data and used to train the models for comparison with non-contact sensors. The results of this study can be used for transfer learning by using various contact sensor data that have proven to be adequate for abnormality detection and comparing them with non-contact sensor data. The proposed approach will also be of use to healthcare professionals who need to monitor various other variables to detect abnormal situations such as arrhythmia. Extensive research on the measurement of biomedical information, such as motion, heart rate, and breath data, using non-contact sensor is anticipated in the near future, and the integration of contact and untact technologies will provide us with considerable ambient intelligence as recommendations for cognitive health advisor platforms [19].

REFERENCES

[1] B. M. Saykrs, “Analysis of heart rate variability,” Ergonomics, vol. 16, no. 1, pp. 17–32, Jan. 1973.
[2] P. Vyas and D. Pandit, “Heartbeat abnormality detection using machine learning models and rate variability (HRV) data,” Preprints, to be published.
[3] W. H. Lin, D. Wu, C. Li, H. Zhang, and Y. T. Zhang, “Comparison of heart rate variability from PPG with that from ECG,” in Proc. Int. Conf. Health Inform., 2014, pp. 213–215.
[4] T. Hall, D. C. Lie, T. Q. Nguyen, J. C. Mayeda, P. E. Lie, J. Lopez, and R. E. Banister, “Non-contact sensor for long-term continuous vital signs monitoring: A review on intelligent phased-array Doppler sensor design,” Sensors, vol. 17, no. 11, p. 2632, 2017.
[5] C. M. Van Ravenswaaij-Arts, L. A. Kollee, J. C. Hopman, G. B. Stoelinga, and H. P. van Geijn, “Heart rate variability,” Ann. Internal Med., vol. 118, no. 6, pp. 436–447, 1993.
[6] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers, and T. D. Wager, “A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health,” Neurosci. Biobehavioral Rev., vol. 36, no. 2, pp. 747–756, Feb. 2012.
[7] J. F. Thayer, S. S. Yamamoto, and J. F. Brosschot, “The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors,” Int. J. Cardiol., vol. 141, no. 2, pp. 122–131, May 2010.
[8] S. S. Guntert and S. Chomal, U.S. Patent 15 147 720, 2016.
[9] S. Kaplan Berkaya, A. K. Uysal, E. Sora Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, “A survey on ECG analysis,” Biomed. Signal. Process. Control, vol. 43, pp. 216–235, May 2018.
[10] A. Droitcour, V. Lubecke, J. Lin, and O. Boric-Lubecke, “A microwave radio for Doppler radar sensing of vital signs,” in IEEE MTT-S Int. Microw. Symp. Dig., vol. 1, May 2001, pp. 175–178.
[11] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek, “Non-contact heart rate and heart rate variability measurements: A review,” Biomed. Signal Process. Control, vol. 13, pp. 102–112, Sep. 2014.
[12] A. J. Cann, M. Malik, J. T. Bigger, G. Breithardt, S. Cerutti, R. J. Cohen, P. Coumel, E. L. Fallen, H. L. Kennedy, R. E. Kleiger, and F. Lombardi, “Heart rate variability: Standards of measurement, physiological interpretation and clinical use,” in Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996.
[13] Z. Öri, G. Monir, J. Weiss, X. Sayhouni, and D. H. Singer, “Heart rate variability: Frequency domain analysis,” Cardiol. Clinics, vol. 10, no. 3, pp. 499–533, 1992.
[14] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” IEEE Eng. Med. Biol. Mag., vol. 20, no. 3, pp. 45–50, May 2001.
[15] D. Praveen Kumar, T. Amgoth, and C. S. R. Annavarapu, “Machine learning algorithms for wireless sensor networks: A survey,” Inf. Fusion, vol. 49, pp. 1–25, Sep. 2019.
[16] Y. Freund and L. Mason, “The alternating decision tree learning algor- ithm,” in Proc. ICML, vol. 99, 1999, pp. 124–133.
[17] T. G. Dietterich, “Ensemble learning,” in The Handbook of Brain Theory and Neural Networks. U.K.: Michael A. Arbib, 2002.
[18] J. A. K. Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” Neural Process. Lett., vol. 9, no. 3, pp. 293–300, Jun. 1999.
[19] K. Lee, “Reference model and architecture of interactive cognitive health advisor based on evolutionary cyber-physical systems,” KSII Trans. Internet Inf. Syst., vol. 13, no. 8, pp. 4270–4284, Aug. 2019.