Hierarchical User Status Classification for Imbalanced Biometric Data Class

Nakyoung Kim
Institute for Information Technology Convergence
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
nkim71@kaist.ac.kr

Hyunseo Park
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
tkf92001@kaist.ac.kr

Gyeong Ho Lee
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
gyeongho@kaist.ac.kr

Jaeseob Han
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
j89449@kaist.ac.kr

Hyeontaek Oh
Institute for Information Technology Convergence
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
hyeontaek@kaist.ac.kr

Jun Kyun Choi
School of Electrical Engineering
Korea Advanced Institute of Science and Technology
Daejeon, Rep. of Korea, 34141
jkchoi59@kaist.ac.kr

***Abstract*—With the proliferation of Internet of Things technologies, health care services that target a household equipped with IoT devices are widely emerging. In the meantime, the number of global single households is expected to rapidly grow. Contactless radar-based sensors are recently investigated as a convenient and practical means to collect biometric data of subjects in single households. In this paper, biometric data collected by contactless radar-based sensors installed in single households of the elderly under uncontrolled environments are analyzed, and a deep learning-based classification model is proposed that estimates a user's status in one of the predefined classes. In particular, the issue of the imbalance class sizes in the generated dataset is managed by reorganizing the classes into a hierarchical structure and designing the architecture for a deep learning-based status classification model. The experimental results verify that the proposed classification model has a noticeable impact in mitigating the issue of imbalanced class sizes as it enhances the classification accuracy of the individual class by up to 65% while improving the overall status classification accuracy by 6%.**

***Index Terms*—Radar-based status monitoring, status classification, imbalanced data**

I. Introduction

With the proliferation of Internet of Things (IoT) technologies, diverse applications and services targeting a household equipped with IoT devices are widely developed. In the meantime, single households are expected to globally grow over the next decades, and the single households of the elderly are anticipated to take a significant portion [1]. Accordingly, the need for health care services for various health problems (e.g., monitoring of underlying diseases, detecting emergencies, etc.) that can occur in single households of the elderly is increasing.

Many health care services utilize on-body or contactless sensors to collect and monitor the data about their service users in their households [2]. The on-body method obtains the data through a wearable device such as a smartwatch, which requires to be attached to the body of the user for service provision. Since the on-body sensors are often tightly placed on the user's body, the biometric data collected from the on-body sensors are relatively accurate. However, the user has to continuously wear the monitoring device for reliable service provision, and the devices need to be periodically charged. This increases the inconvenience and leads to a low usage rate, in particular, for the high age group who often has difficulties in using smart devices [3]. Accordingly, the methods to monitor the users' status in their households with externally installed contactless sensors are widely investigated for reliable service provision without affecting the users' daily life [4].

With the advances of the sensor technology, radar-based data acquisition methods are recently investigated, which obtains an individual's biometric information (e.g., heart rate, respiration, etc.) from the phase change of the radar reflected by the movement of the human body, and the development of relevant data processing methods are in progress [5]–[8]. In this paper, a classification method is studied to estimate the status of subjects from their real-world biometric information

This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE) and Korea Institute for Advancement of Technology (KIAT) through the International Cooperative R&D program. (Project No. 0011879)

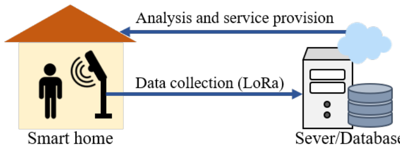

Fig. 1. High level radar-based biometric monitoring system model

Fig. 2. Actual illustration of the radar sensor installed in an single household

TABLE I
SPECIFICATIONS OF THE INSTALLED RADAR SENSOR

Category	Specification
Chip	Sharp DC6M4JN3000
Method	Microwaves (24.05 24.25GHz)
Resolution	60cm
Range (Max.)	1.5m (Heartbeats, Breathing) 7m (Body motion)
Directionality	Azimuth: 25°, Elevation: 20°
Error rate	±10% (3m)

that is collected through contactless radar-based sensors installed in single households. As the frequency and duration for the occurrence of the subjects' status differ, the issue of status class imbalance inevitably rises in uncontrolled environments. In this paper, the status classes in the collected data are reorganized into a hierarchical structure to handle the issue of the imbalanced classes. In particular, this paper analyzes the biometric information of 22 elders collected in uncontrolled environments and proposes a subject's status classification model with the hierarchically structured data that alleviates the imbalance issue.

The rest of this paper is organized as follows. Section II provides previous studies. Section III introduces the status classification method. The performance evaluation is presented in Section IV, followed by the conclusion in Section V.

II. Related Works

Biometric information monitoring technologies with contactless radar-based sensors are largely divided into two fields of research: research on increasing the accuracy of the measurement of monitoring devices and research on accurately and effectively analyzing the obtained information. As the sensor technologies have become more mature than before, the studies to accurately analyze the obtained data are actively in progress nowadays. In relation to contactless radar-based biometric information monitoring, research on cardio-respiration monitoring is widely investigated, which is mainly conducted in an impulse-radio ultra wide band (IR-UWB) radar-based environment [5]. These studies include the research on accurate measuring of heart rate, continuous monitoring of cardio-respiration rate [9], and monitoring and estimating the sleep stages [8]–[10]. In particular, with the development of machine learning and artificial intelligence technologies, various studies have now adapted deep learning technologies to accurately analyze biometric data. In [10], the cardio-respiratory signals collected through an IR-UWB radar-based device are analyzed with recurrent neural networks and an attention mechanism. In [11], the performance of various deep learning technologies with the publicly available biometric datasets is evaluated. In [12] and [13], electrocardiogram signals are analyzed based on a convolutional neural network.

However, to the best knowledge of the authors, the previous studies have not managed the issue of the imbalance classes in the dataset that inevitably occurs in practice under uncontrolled environments. Accordingly, this study specifically targets a method to alleviate the adverse effect caused by the imbalanced dataset that is used to develop a user's status classification model with the biometric data collected by a contactless radar-based sensor.

TABLE II
SAMPLE DISTRIBUTION OF THE HR DATASET

Class	Status	Number of data samples	Ratio
1	Not detected	3,673	34.85%
2	In sleep	4,831	45.84%
3	In active activity	602	5.71%
4	In stationary Activity	554	5.26%
5	In unidentified activity	879	8.34%
Total		10,539	100%

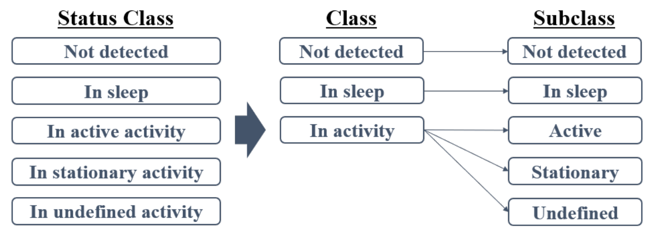

Fig. 3. Reorganization of the status dataset structure with biometric data

III. Status Classification from Imbalanced Data of Biometric Feature

In this section, we propose a deep learning-based status classification model based on the biometric data of a single targeted subject. The proposed model estimates the status of the subject in one of the predefined status classes. In prior to providing the details of the proposed model, we introduce the dataset first to ease the understanding of the reason behind the proposed classification model.

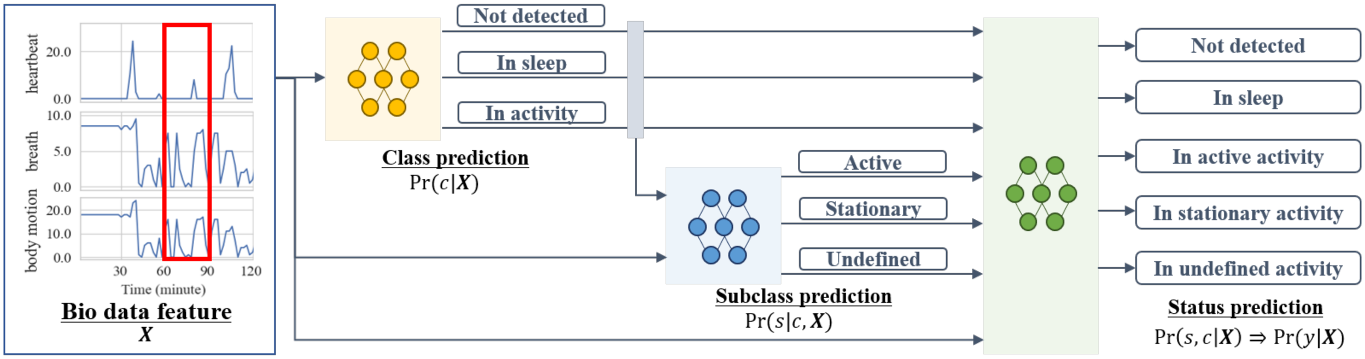

Fig. 4. Proposed deep learning-based user status estimation model with biometric data with imbalanced classes

A. Dataset Generation and Data Preprocessing

The real-world biometric data that were collected from uncontrolled environments are used in this study. The radar sensors with the specifications given in Table I were installed in a bed room of over 100 individual single households, as described in Fig. 1 and Fig. 2. The dataset was generated by collecting the heart rate, respiratory rate, and body motion, and other types of biometric features for every 2 minutes at different times for each individual from 2020 to 2021. Whereas the dataset includes various types of biometric features, the heart rate, respiratory rate, and body motion data are used in this paper for simplicity. However, the method proposed in this paper is not restricted to the specific combination of the biometric features but can be extended to many feature compositions.

The subject's status for the collected biometric features are labeled by experts as one of the five predefined classes: the subject is 1) not detected, 2) in sleep, 3) in an active activity, 4) in a stationary activity, and 5) in an undefined activity. From the generated dataset, the parts that are stably collected without a network disconnection are extracted, resulting in a subset of the data from 22 out of 112 subjects. Min-max normalization is applied to each dimension of the biometric features. The data are segmented into fully non-overlapping data samples for 30 minutes, and the dimension of a data sample accordingly becomes 15. The status of a segmented data sample is set as the most frequently detected status in that period. Table. II shows the sample distributions of the dataset with respect to the status classes.

As the dataset is collected with externally installed sensors in the subjects' places, the dataset includes the data when the subjects are in sleep as well as when they are out of the range of the sensors. Subsequently, the frequency and duration of the subjects' statuses are different under uncontrolled environments, and the imbalance in the status classes inevitably occurs. In the remainder of this section, the issue of the imbalanced classes is handled by reorganizing the status classes into a hierarchical structure with classes and subclasses as described in Fig. 3 and by developing a neural network model that is particularly designed to learn the hierarchical structure.

B. Status Classification

The objective of a status classification problem is to estimate the probability of occurrence of a status in a given period when a sequence of biometric features is observed, such that

$$\Pr(y \mid \boldsymbol{X}), \tag{1}$$

where y denotes the status class before the reorganization, and $\boldsymbol{X}$ is the input sequence of biometric features. As the dataset is reorganized in a hierarchical structure, the status classification problem in (1) can be now considered as estimating the joint probability of classes and subclasses, which is given by

$$\Pr(s, c \mid \boldsymbol{X}), \tag{2}$$

where c denotes the class, and s denotes the subclass. The joint probability in (2) then can be further divided into

$$\Pr(c \mid \boldsymbol{X}) \Pr(s \mid c, \boldsymbol{X}). \tag{3}$$

where $\Pr(c \mid \boldsymbol{X})$ indicates the probability of the occurrence of a class when a biometric data feature is observed, and $\Pr(s \mid c, \boldsymbol{X})$ indicates the probability of the occurrence of a subclass when the information on the class and biometric data are given.

That is, the status classification problem in (1) can be investigated as a combination of two separate classification problems, which are respectively predicting the class and subclass. Through separately developing these classification models, the problem to find the decision boundaries among the classes and subclasses can be simplified. In addition, the issue of imbalance status class can be mitigated to some degree by transferring the classification models that are separately pre-trained into a comprehensive model for integration and parameter tuning as described in Fig. 4.

As the main focus of this paper is to investigate the impact of the proposed classification model structure particularly designed for the hierarchically structured dataset, multi-layered perceptrons (MLP) are applied for the classification models for simplicity. However, the proposed model is not restricted to MLP but other advanced deep learning models can be alternatively used to enhance the performance.

TABLE III
CONFUSION MATRIX FOR ERROR MEASUREMENTS

Actual \ Predicted	Positive	Negative
Positive	True positive (TP)	False negative (FN)
Negative	False positive (FP)	True negative (FN)

IV. PERFORMANCE EVALUATION

In this section, the performance of the proposed status classification model is evaluated in terms of its classification accuracy.

A. Evaluation Metric

The performance of the proposed model is evaluated with a well-known error measurement, F_1 score, based on the confusion matrix illustrated in Table III. The F_1 score is the harmonic mean of the precision and recall, where the precision indicates the ratio between the true positive instances and all positive instances while the recall indicates the ratio between the true positive instances and all of the instances that are identified as positive. For a multi-label classification problem, the confusion matrix and F_1 score of each class are computed and combined afterward.

The F_1 score for the i-th class is given by

$$\mathrm{F}_1 \text{ score} = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \frac{1}{2}\left(\mathrm{FP}_i + \mathrm{FN}_i\right)}, \tag{4}$$

where TP_i, TN_i, FP_i, and FN_i are the numbers of the true positive, true negative, false positive, and false negative instances for the i-th class. The F_1 score has a range of $[0, 1]$, and a value closer to 1 implies better classification performance. The unweighted mean of each class' F_1 scores is widely considered an adequate metric to evaluate the performance of a classifier on a dataset with imbalanced class samples. Accordingly, the equally weighted F_1 scores of the classes regardless of their numbers of instances are used for performance evaluation.

B. Experimental Settings and Results

To evaluate the performance of the proposed classification model, experiments on the classification methods with conventional machine learning and deep learning techniques are conducted. As the benchmarks, the performance of Gaussian naïve Bayes (NB), k-nearest neighbor (kNN), support vector machine (SVM), and random forest (RF) are experimentally investigated. In addition, MLP with different numbers of layers and nodes are used as the baselines. The conventional machine learning methods are implemented with the scikit-learn library in python. Each part of the proposed model is designed with two fully connected layers where each layer is respectively composed of 16 and 8 nodes. The class and subclass classification models are pre-trained and combined with the status classification part in the end for further training. The baseline MLP models are designed with a combination of layers with 32, 16, and 8 nodes. The deep learning models are implemented with TensorFlow and Keras. For all deep learning models, Adam optimizer is applied with a learning rate 0.001, and they are trained until the losses are converged with the learning rate decay and early stopping. For training, validation, and test, 60%, 20%, and 20% of instances are randomly selected from the generated dataset, respectively.

TABLE IV
F_1 SCORE OF THE STATUS CLASSES (%)
(STATUS CLASS: 1. NOT DETECTED, 2. IN SLEEP, 3. IN ACTIVE ACTIVITY, 4. IN STATIONARY ACTIVITY, AND 5. IN UNDEFINED ACTIVITY)

Model	Status class 1	2	3	4	5	Avg.
NB	**83.3**	**74.6**	**39.0**	**18.9**	**63.2**	**55.8**
kNN	88.1	84.1	11.9	12.9	61.1	51.6
SVM	90.2	88.6	7.6	3.7	72.8	52.6
RF	92.3	87.4	5.5	15.9	73.3	54.9
MLP (16-8)	90.7	91.3	53.2	13.1	80.9	65.8
MLP (16-16-8)	**91.0**	**90.9**	**44.2**	**24.2**	**78.9**	**65.8**
MLP (16-16-16-8)	90.6	90.6	46.4	5.3	79.0	62.4
MLP (32-16-8)	91.0	91.6	49.3	11.9	80.5	64.9
MLP (32-16-16-8)	91.0	90.8	50.0	9.8	80.7	64.5
Proposed	**91.3**	**91.0**	**46.1**	**40.2**	**80.9**	**69.9**

The experimental results are provided in Table IV. The results show that the deep learning approaches outperform the conventional machine learning approaches. In addition, as the number of instances in Class 1 (not detected) and Class 2 (in sleep) are much larger than those of the other classes, the classification models are trained biased to Class 1 and Class 2. As result, the F_1 scores of the classes with small numbers of instances, Class 3 (in an active activity) and Class 4 (in a stationary activity), are fairly lower than the others in general. In particular, Class 4 is relatively difficult to be separated from Class 2 as the biometric features of these two classes are not much different from each other but Class 4 has much fewer instances compared to Class 2. Even though the MLP models achieve higher F_1 scores for Class 3 than the conventional machine learning models do, their F_1 scores for Class 4 are still comparatively low. In the meantime, the proposed model achieves noticeable improvements on the F_1 scores of both Class 3 and Class 4, resulting in the highest F_1 score on average over the classes.

In summary, whereas there have not been significant improvements in the classification accuracy for the classes with a large number of instances, the results show that the proposed model improves the classification accuracy for the classes in small sizes. Hence, these results verify the effectiveness of the proposed classification model in the structure that is particularly designed for a dataset with imbalanced class sizes.

V. CONCLUSION

In this paper, the biometric data (i.e., heart rate, respiratory rate, and volume of motion) collected in single households of the elderly under uncontrolled environments are investigated, and a classification model is proposed that estimates a user's status in one of the five predefined classes (i.e., not detected, in sleep, in an active activity, in a stationary activity, and in an undefined activity). In particular, the issue of the imbalanced class is managed by reorganizing the classes into a hierarchical structure and separating the status classification problem into two disjoint problems based on probabilistic analysis. The separated classification models are deliberately designed to manage the individual objective of the disjoint classification problems. The scope of this study is to verify the impact of the structure of the proposed classification model in alleviating the adverse effects of the imbalanced class. Accordingly, simple feed-forward networks are applied for classification to rule out the influences caused by the choice of deep learning technologies other than the structure of the classification model itself. The experimental results show that the proposed model enhances the status classification accuracy of the individual class by over 65% while improving the overall accuracy by 6%. Whereas the improvements in the overall accuracy are not significant, the results verify that the proposed model has a noticeable impact in managing the issue of imbalanced class. To improve the overall classification accuracy, more advanced deep learning technologies (e.g. convolutional neural network, long-short-term memory, etc.) can be extensively applied to the proposed classification model structure as a future work of this study.

ACKNOWLEDGMENT

The authors of this paper thank Seung Chul Kim from KULS, João Garcia from Ubiwhere, and Ricardo Gonçalves from PROEF for co-working to develop the dataset and prototype for the proposed system as partners of a research project (KIAT, Project No. 0011879).

REFERENCES

[1] Organisation for Economic Co-operation and Development (OECD). The Future of Families to 2030. Accessed: Feb., 2020. [Online]. Available: https://doi.org/10.1787/9789264168367-en

[2] H. Mshali, T. Lemlouma, M. Moloney, and D. Magoni, "A survey on health monitoring systems for health smart homes," *International Journal of Industrial Ergonomics*, vol. 66, pp. 26–56, 2018.

[3] J. Kim, "How much do we know about the use of smartphones in the silver generation?: Determinants of the digital divide within the silver generation," *Information Society & Media*, vol. 21, no. 3, pp. 33–64, 2020.

[4] M. Bahache, J. P. Lemayian, W. Wang, and J. Hamareh, "An inclusive survey of contactless wireless sensing: A technology used for remotely monitoring vital signs has the potential to combating covid-19," 2020.

[5] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek, "Non-contact heart rate and heart rate variability measurements: A review," *Biomedical Signal Processing and Control*, vol. 13, pp. 102–112, 2014.

[6] A. Ni, A. Azarang, and N. Kehtarnavaz, "A Review of Deep Learning-Based Contactless Heart Rate Measurement Methods," *Sensors*, vol. 21, no. 11, 2021.

[7] F. Zhang, C. Wu, B. Wang, M. Wu, D. Bugos, H. Zhang, and K. J. R. Liu, "Smars: Sleep monitoring via ambient radio signals," *IEEE Transactions on Mobile Computing*, vol. 20, no. 1, pp. 217–231, 2021.

[8] Y. Lee, J. Y. Park, Y. W. Choi, H. K. Park, S. H. Cho, S. H. Cho, and Y. H. Lim, "A Novel Non-contact Heart Rate Monitor Using Impulse-Radio Ultra-Wideband (IR-UWB) Radar Technology," *Scientific Reports*, vol. 8, no. 13053, 2018.

[9] C. H. Chang and W. W. Hu, "Design and implementation of an embedded cardiorespiratory monitoring system for wheelchair users," *IEEE Embedded Systems Letters*, vol. 13, no. 4, pp. 150–153, 2021.

[10] H. B. Kwon, S. H. Choi, D. Lee, D. Son, H. Yoon, M. H. Lee, Y. J. Lee, and K. S. Park, "Attention-based lstm for non-contact sleep stage classification using ir-uwb radar," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 10, pp. 3844–3853, 2021.

[11] S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmarking deep learning models on large healthcare datasets," *Journal of Biomedical Informatics*, vol. 83, pp. 112–134, 2018.

[12] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "Ecg heartbeat classification: A deep transferable representation," in *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, 2018, pp. 443–444.

[13] N. Ahmed, A. Singh, S. KS, G. Kumar, G. Parchani, and V. Saran, "Classification Of Sleep-Wake State In A Ballistocardiogram System Based On Deep Learning," 2020.