Nurse Care Activity Recognition Challenge: A Comparative Verification of Multiple Preprocessing Approaches

Hitoshi Matsuyama
hitoshi@ucl.nuee.nagoya-u.ac.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

Takuto Yoshida
takuto@ucl.nuee.nagoya-u.ac.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

Nozomi Hayashida
linda@ucl.nuee.nagoya-u.ac.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

Yuto Fukushima
fukurin@ucl.nuee.nagoya-u.ac.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

Takuro Yonezawa
takuro@nagoya-u.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

Nobuo Kawaguchi
kawaguti@nagoya-u.jp
Graduate School of Engineering,
Nagoya Uniersity
Aichi, Japan

ABSTRACT

Although activity recognition has been studied considerably for the last two decades, it is still not so easy to handle complicated activity classes in a specific domain. The 2nd Nurse Care Activity Recognition Challenge Using Lab and Field Data aims to explore a part of those complicated activities by focusing on the nurse caring. Our team, "UCLab", found that the main problem in the challenge is the imbalance and unevenness of the dataset, each of which often happens in real-field data. Considering the problem, we approached the challenge using a Random Forest-based method with multiple preprocessing to classify 12 activity modes. Our approach consists of the following steps: We first preprocessed the acceleration data to obtain uniformly sampled signals. Then we extracted acceleration data with respect to each row of the given label data and extracted feature values. We adopted Random Forest for classification and performed post-processing to the predicted data obtained from the classifier. As a result, we obtained 51.5% accuracy with the trial-based evaluation.

CCS CONCEPTS

• **Applied computing** → *Health care information systems.*

KEYWORDS

datasets; activity recognition; signal processing; nurse care

ACM Reference Format:

Hitoshi Matsuyama, Takuto Yoshida, Nozomi Hayashida, Yuto Fukushima, Takuro Yonezawa, and Nobuo Kawaguchi. 2020. Nurse Care Activity Recognition Challenge: A Comparative Verification of Multiple Preprocessing Approaches. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers (UbiComp/ISWC '20 Adjunct), September 12–16, 2020, Virtual Event, Mexico.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3410530.3414333

UbiComp/ISWC '20 Adjunct, September 12–16, 2020, Virtual Event, Mexico

ACM ISBN 978-1-4503-8076-8/20/09…$15.00
https://doi.org/10.1145/3410530.3414333

1 INTRODUCTION

Over the last two decades or more, Human Activity Recognition (HAR) has emerged as a considerably important research topic and many researchers have been developed. While many researchers have investigated vision-based HAR [1], inertial sensor-based HAR has also emerged as an important research topic [2].

HAR using inertial sensors has advanced with the growth of sensor technology. These days, its potential is extended to various applications, such as industry [3], office[4], sports[5], and nurse caring scenarios[6]. In particular, nurse caring is becoming an extremely important target with the rapid growth of the elderly population in the world.

Although HAR using inertial sensors has been studied considerably, it is still not so easy to handle complicated activity classes in a specific domain. One of the reasons for that is the lack of data on those activities. To promote investigation in those activities, The 2nd Nurse Care Activity Recognition Challenge Using Lab and Field Data [7] [8] opens a part of the Nurse Activity Dataset which includes 12 different activities collected from 8 participants. The purpose of this challenge is to build an activity classification model for testing data which includes acceleration data collected from 3 participants in real fields.

Our team, "UCLab", approaches to the challenge with the following steps: We first preprocessed the acceleration data to obtain uniformly sampled signals. Then we extracted acceleration data with respect to each row of the given label and extracted feature values. We adopted Random Forest for classification, and then performed post-processing to the predicted data obtained from the classifier. As a result, we obtained 51.5% accuracy with the trial-based evaluation.

2 DATASET DESCRIPTION

The dataset [7] contains nurse activity data collected from Accelerometer in the smartphone which is attached on the right arm using the armband. There are 12 types of activities, each of which are done in the Care facility. The activities can be categorized into 3 principal modes:

- Help in Mobility
- Assistance in Transfer
- Position Change

and named as Category A, B, and C respectively. Category A has 4 modes of activities, the names of which are Guide(from the front), Partial assistance, Walker, and Wheelchair. Category B has 4 modes of activities, named as All assistance, Partial assistance (From the front), Partial assistance (From the side), and Partial assistance (From the back) respectively. Category C also has 4 modes of activities, however, two of them are categorized in the same label. They are To supine position/To Right lying position, To Left lying position, Lower body lifting, and Horizontal movement.

The data were collected both in an experiment lab and in the real field. Training data consists of two types: lab data and field data. The lab data was collected in the Smart Life Care Unit of the Kyusyu Institute of Technology in Japan. It contains acceleration data from 2 professional nurses. The field data is collected in a Care Facility in Japan. It contains data of 6 nurses. Acceleration data is collected using mobile phones attached in the right arm using the armband. The sampling rate is 60 Hz and there no preprocessing applied to the data. On the other hand, testing data contains acceleration data of 3 participants in the real field. In contrast to the training data, there are no data collected in the lab.

3 METHODOLOGY

Figure 1 shows an overview of our approach. In order to work out our strategy, we first investigated the dataset. The first thing we noticed is that compared to lab data, which is well-balanced among all activity types, field data contains very imbalanced activity classes. Figure 2 shows the duration of each activity class in field data. Compared to activity modes of 3 and 5, other activity modes contain fewer durations. We also noticed that in some parts of the training data, the sampling rate gets lower compared to the other points. Thus, our approach contains some preprocessing. After that, we extract feature values to recognize the characteristics of each activity.

3.1 Preprocessing

Our team applied a resampling algorithm and sliding window method to fit our machine learning model.

For the training data, we first extracted the corresponding acceleration data from each row of the given label data, and then removed the sparse data with less than six samples per unit time (which means that the average sampling frequency is less than 10 Hz), as it was considered difficult to extract features of those waveforms. Fortunately, the test data contained a lot of dense data compared to the training data. Therefore, we expect it is possible to improve the accuracy of testing by removing the sparse data from the training data. Furthermore, to decrease the bias in the training data, we removed data that contains more than 50,000 samples in a single row of label data.

We then resampled the extracted data at 10 Hz using cubic spline interpolation to create a sliding window with a window width of 8 seconds and a 90% overlap. Sampling frequency was reduced to 10

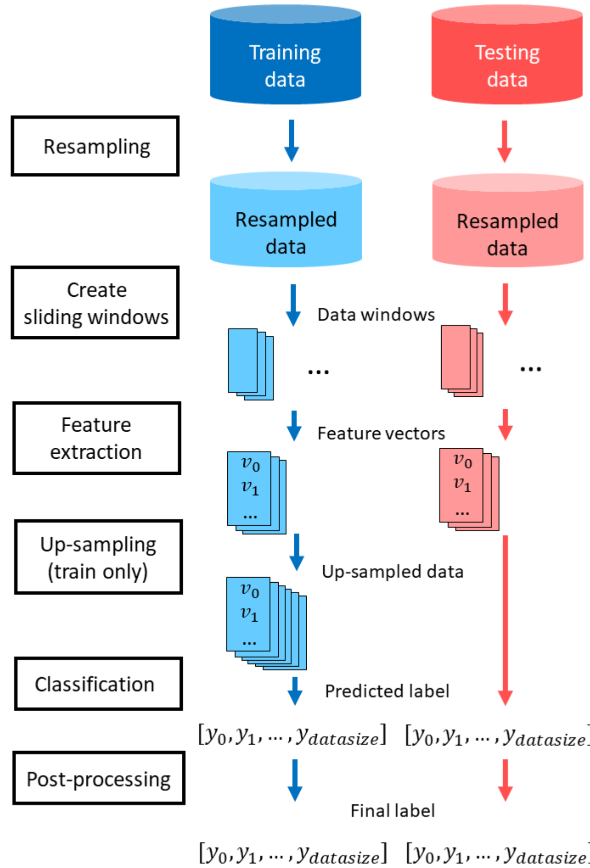

Figure 1: Method Overview

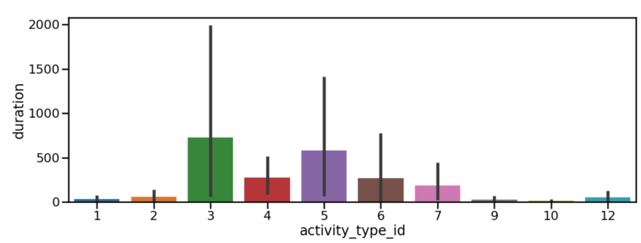

Figure 2: Duration(sec.) of Each Activity Class in Field Data

Hz because the sampling frequency of the original data was smaller than 10 Hz on average.

In the 2nd Nurse care activity recognition challenge dataset, there is little sensor data for which the corresponding labels exist. Furthermore, neither the amount of data per activity mode nor the amount of data per user is well-balanced. Therefore, it is not appropriate to apply the basic cross-validation or the leave-one-subject-out method to the dataset because neither of them can split the data equally. Therefore, we developed a trial-based cross-validation method to evaluate the accuracy of the dataset properly. Figure 3 shows how our developed cross-validation method works. The splitting method is to split data based on the index of the label

Figure 3: Trial-based Train-test Split Overviews

data. Here, each label is split uniformly. Although the accuracy of this split validation data is smaller than that of the usual method, we consider that this approach is the most proper way to evaluate the accuracy of the dataset because it does not allow the same data to appear in both training and validation data and does not cause overlearning. Furthermore, by splitting the data in this way, even if the overlap of the sliding window is large, overlearning will not occur. Therefore, we increased the amount of training data by extending the overlap to 90%. Finally, the SMOTE (Synthetic Minority Oversampling Technique [9]) algorithm was used to oversample the data in order to ensure that the number of data in each activity mode in the training data is equal.

For the test data, we first divided the acceleration data into intervals with timestamps more than 10 seconds apart. This is because we judged that if there was no data for more than 10 seconds, we can consider that the behavior was switched. We then resampled the segmented data at 10 Hz using cubic spline interpolation to create a sliding window with a window width of 8 seconds and an overlap rate of 90%, which is the same format as the training data.

3.2 Feature Extraction

In this section, we will provide an overview of our feature extraction method. The activity modes in this challenge consist of constant behaviors such as "help in mobility" and temporary activities such as "assistance in transfer" and "position change". Mean and variance of waves are generally used for constant activities such as walking. On the other hand, since the waveform itself is often used for temporary activities, we extracted the basic features that express the waveform: percentile, minimum, maximum, kurtosis, skewness, and signal power. We also extracted the norm mean, variance, and median to extract activity features that are independent of the sensor orientation. Including these features, we used 35 types of features in total.

The sampled sensor data is denoted by x_i, y_i, z_i and the data sequence consists of N samples is denoted by $X = \{[x_i, y_i, z_i] | i = 1, 2, ..., n\}$.

The mean of x ($\bar{x}$) is represented by the following formula(1).

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{n} x_i \tag{1}$$

The variance value is a measure of how far a data sequence X is from its mean value. The variance of x (σ^2) is represented by the following formula(2).

$$\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{2}$$

The 25th percentile is known as the first quartile, the 50th percentile as the median or second quartile, and the 75th percentile as the third quartile of the total when data column X is arranged in decreasing order.

The minimum of x (x_{min}) and the maximum of x (x_{max}) is represented by the following formulas(3)(4).

$$x_{min} = min\{x_i | i = 1, 2, ..., n\} \tag{3}$$

$$x_{max} = max\{x_i | i = 1, 2, ..., n\} \tag{4}$$

The skewness expresses the asymmetry of the distribution of the data sequence X. The skewness of x (β_{1x}) is represented by the following formula(5).

$$\beta_{1x} = \frac{1}{N}\sum_{i=1}^{n}(\frac{x_i - \bar{x}}{\sigma_x})^3 \tag{5}$$

The kurtosis expresses the sharpness of the distribution of the data sequence X. The kurtosis of x (β_{2x}) is represented by the following formula(6).

$$\beta_{2x} = \frac{1}{N}\sum_{i=1}^{n}(\frac{x_i - \bar{x}}{\sigma_x})^4 \tag{6}$$

The signal power of x ($\bar{x^2}$) is represented by the following formula(7).

$$\bar{x^2} = \frac{1}{N}\sum_{i=1}^{n} x_i^2 \tag{7}$$

The covariance is a measure of the relationship between two types of data. The covariance of x and y (S_{xy}) is represented by the following formula(8).

$$S_{xy} = \frac{1}{N}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{8}$$

3.3 Classification Model

As a classification model, we adopted the Random Forest algorithm [10]. Compared to other classifiers like Support Vector Machine, it works better in multi-class classification problems. In addition, it has an advantage in accuracy with relatively small datasets compared to Deep Learning-based methods. The summary of the last year's Nurse Care Challenge [11] indicates a simple classification approach by Md. Eusha Kadir et al. [12] worked better compared to the other approaches. Although the approach by Md. Eusha Kadir et al. uses the KNN classifier, we consider Random Forest suits better in this challenge. This is because the data size of some activity modes in the training data is small. As the KNN is a model of low-bias and high-variance, it sometimes overfits the training data and causes low accuracy when the data size is relatively small [13]. Thus, we

Table 1: Evaluation Result

	base line	trial1	trial2	trial3	trial4
P1	✓	✓	✓	✓	✓
P2		✓	✓	✓	✓
P3			✓	✓	
P4				✓	✓
accuracy [%]	40.9	47.8	40.6	48.6	**51.5**

adopt Random Forest, which is a model using ensemble learning and may handle that bias-variance problem.

3.4 Post-processing

This section describes the specific methods of post-processing. Our method is performed in two steps. First, the majority voting was applied to each timestamp using multiple windows of the predicted labels obtained from the classifier. After that, if the percentage of the most common result was not more than 60%, a new majority voting was applied by combining the results from the previous and subsequent timestamps, up to a maximum of 8 (approximately the width of the window).

4 VALIDATION RESULT

In this chapter, we experiment to find the preprocessing combination that gives the best validation accuracy. The four candidates for preprocessing is as follows:

- P1 : resampling (10Hz)
- P2 : removing sparse data
- P3 : over-sampling (SMOTE)
- P4 : removing too long records

P1 is resampling the segmented data at 10 Hz using cubic spline interpolation. P2 is removing the sparse data which is difficult for models to learn features. P3 is used to over-sample the data in order to ensure that the number of data in each activity mode in the training data is equal. P4 is removing a single row of label data in which the duration of the recording is too long.

These preprocessing methods are combined and applied to the data before performing classification with Random Forest to estimate and obtain the validation result. Table 1 shows the experimental results. We define the result with P1 as the baseline accuracy. In trial 1, the accuracy is improved over the baseline. The result shows that P2 contributes to improving the accuracy by approximately 7%. On the other hand, trial2 is less accurate than trial1 by adding P3 to the preprocessing. In trial3, we used all four preprocessing methods and achieved a 48.6% accuracy. In trial4, we used P1, P2, and P4, and achieved a 51.5% accuracy, exceeding the accuracy of all other trials.

These results indicate that P2 and P4 are effective in improving the accuracy of the classification. In particular, P2 made a significant contribution to improve accuracy. This implies that sparse data is a major hindrance to model learning. Figure 4 shows the confusion matrix of our method with all preprocessing.

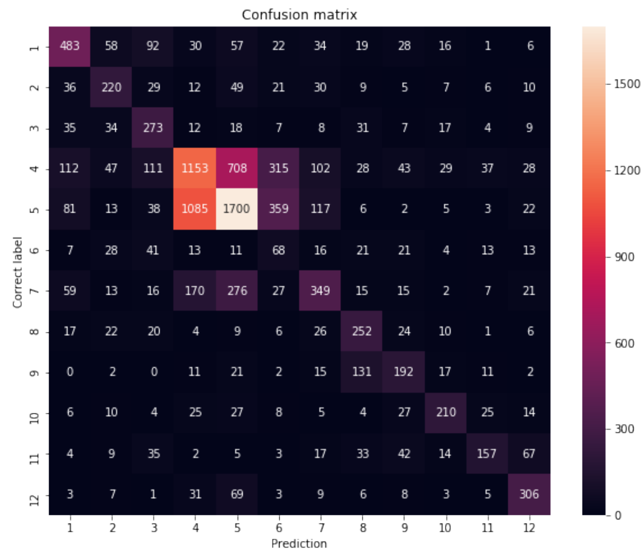

Figure 4: Confusion Matrix of the Original Labels v.s. Predicted Labels with All Preprocessing

5 CONCLUSION

In this paper, we presented a Random Forest-based method while comparing multiple preprocessing approaches for the Second Nurse Care Activity Recognition Challenge. The main problem we found in this challenge is the imbalance and unevenness of the dataset, each of which often happens in real-field data. To handle these problems, we applied four preprocessing methods. First, we removed the sparse data whose sampling rate are less than 10 Hz. After that, we applied cubic spline interpolation to the data to obtain uniformly sampled signals. Then we removed a single row of label data. Finally, we applied SMOTE to training data to handle the imbalance of field data and obtain an evenly balanced data among all activity modes. We adopted Random Forest for classification and performed post-processing to the predicted data obtained from the classifier. As a result, we obtained 48.6% accuracy with the trial-based evaluation applying all preprocessing, and 51.5 % accuracy applying all preprocessing except SMOTE. Although the accuracy with SMOTE is relatively lower, we consider this is due to the imbalance of validation data. As it is not possible to know the balance among activity modes of testing data, we applied all of the four preprocessing methods to the testing data and obtained our final result.

The most important problem we could not have solved is to recognize the activity modes that are too similar in their waveforms. Figure 5 is an example of such waveforms. We can see that there are four different activities connected and show almost no difference in their waveforms. We can see that the activity modes in the figure are the ones that are found to be difficult to be distinguished in Figure 4. Despite the effort in preprocessing and modeling, we could not find out a method to distinguish them. In the future, the accuracy may be improved by using the information other than the data waveforms themselves, such as motion transitions between

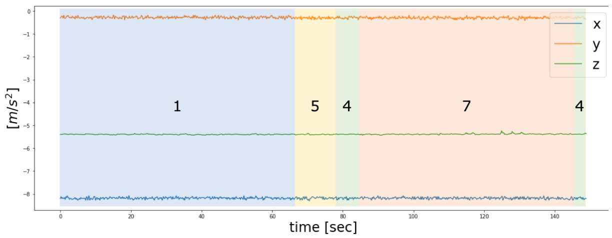

Figure 5: Comparison of 3-axis Acceleration for Label 1, 4, 5 and 7

Table 2: Summary of Resources and Other Information

Used sensor modalities	x, y, and z axis of acceleration
Features used	Described in Section 3
Programming language and libraries	Python3, Numpy, Pandas, Seaborn, Scikit-learn
Window size and Post processing	8 seconds, Described in 3.4
Training and testing time	Training: 92 seconds, Testing: 91 seconds
Machine specification	RAM: 64GB, CPU: Threadripper 1950X(3.5GHz)

activity modes. The recognition result for the testing dataset will be presented in the summary paper of the challenge [8].

REFERENCES

[1] Robert Bodor, Bennett Jackson, and Nikolaos Papanikolopoulos. Vision-based human tracking and activity recognition. In *Proc. of the 11th Mediterranean Conf. on Control and Automation*, volume 1, 2003.

[2] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput. Surv.*, 46(3), January 2014.

[3] I. Maurtua, P. T. Kirisci, T. Stiefmeier, M. L. Sbodio, and H. Witt. A wearable computing prototype for supporting training activities in automotive production. In *4th International Forum on Applied Wearable Computing 2007*, pages 1–12, 2007.

[4] C. Wojek, K. Nickel, and R. Stiefelhagen. Activity recognition and room-level tracking in an office environment. In *2006 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 25–30, 2006.

[5] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. In *23th International Conference on Architecture of Computing Systems 2010*, pages 1–10, 2010.

[6] D. Sánchez, M. Tentori, and J. Favela. Activity recognition for the smart hospital. *IEEE Intelligent Systems*, 23(2):50–57, 2008.

[7] Second nurse care activity recognition challenge. https://abc-research.github.io/nurse2020/. Online, Last accessed: 2020-07-09.

[8] Sayeda Shamma Alia, Paula Lago, Kohei Adachi, Tahera Hossain, Hiroki Goto, Tsuyoshi Okita, and Sozo Inoue. Summary of the 2nd nurse care activity recognition challenge using lab and field data. Proc. HASCA 2020.

[9] Kevin W. Bowyer, Nitesh V. Chawla, Lawrence O. Hall, and W. Philip Kegelmeyer. SMOTE: synthetic minority over-sampling technique. *CoRR*, abs/1106.1813, 2011.

[10] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.

[11] Paula Lago, Sayeda Shamma Alia, Shingo Takeda, Tittaya Mairittha, Nattaya Mairittha, Farina Faiz, Yusuke Nishimura, Kohei Adachi, Tsuyoshi Okita, François Charpillet, and Sozo Inoue. Nurse care activity recognition challenge: Summary and results. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '19 Adjunct, page 746–751, New York, NY, USA, 2019. Association for Computing Machinery.

[12] Md. Eusha Kadir, Pritom Saha Akash, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. Can a simple approach identify complex nurse care activity? In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '19 Adjunct, page 736–740, New York, NY, USA, 2019. Association for Computing Machinery.

[13] Phan Thanh Noi and Martin Kappas. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors (Basel, Switzerland)*, 18(1):18, 12 2017.