Automatic Evaluation System for Nurses' Patient Transfer Skills by Using Deep Learning

Zhihang ZHONG, M1
Ota Laboratory
zhong@race.t.u-tokyo.ac.jp

Abstract—This paper explored the performance of Deep Learning method (Deep-RNNs) for special HAR (Human Activity Recoginiton) task. Compared with human activity of daily living, the attribution of professional nursing activity is very different. Proposed system showed that Deep-RNNs can work well for recognition of separate action. However, as for continuous professional nursing actions, conventional sliding windows based system is not robust enough for continuous professional nursing actions.

Index Terms—Deep Learning, RNN, HAR, Nursing Care

I. Introduction

Aging population causes a greater need of nursing cares [1], but the number of qualified nurses from school cannot meet the requires of society [2]. The ratio of nursing teachers to nursing students is a main reason for this problem. Due to the limited number of nursing teacher, nursing students are difficult to obtain enough practice experience and feedback from professonal teachers [3]. It's hard for them to master the nursing care skills after graduation, which may cause them to make mistakes in actual work environment and get related occupational diseases. Therefore, an evaluation system by which nurse trainee can achieve selflearning becomes an urgent need. For such evaluation system, how to realize the recognition of professional actions, including the right way and wrong way for a same action step, will be the main problem.

II. Related Works

Recent years, the research field of HAR (human activity recognition) obtain a great attention from researchers. Due to the development of deep learning, deep learning methods such as CNN, RNN started to be used in this field, which can realize high accuracy without hand-crafted features compared with classical machine learning methods.

CNN (Convolutional Neural Networks) has a huge impact around 2014, which achieved remarkable results in computer vision, natural language processing, and speech recognition. But it had not been exploited in the field of HAR. At that time, in order to improve the accuracy, most existing work in HAR rely on heuristic hand-crafted feature design and classical machine learning classifications, such as DT (Decision Tree), RF (Random Forest), NB (Naïve Bayes), SVM (Support Vector Machine), KNN (K-Nearest Neighbors).

Therefore, in 2015, [4] [5] built the CNN model for HAR task. They achieved high recognition accuracy without exploring hand-crafted features from time-series signals. Then, [6] compared the recognition performance of CNN (Deep learning) and classical machine learning methods, which showed the power of the deep learning to automatically extract relevant features and achieve slightly better performance.

Later, because LSTM (Long Short Term Memory) based RNN (Recurrent Neural Networks) made great success for time-series domain, researchers started to use RNN model for HAR. In 2016, [7] presented a feature descriptor combined with multi-column BiRNN (Bidirectional RNN) to improve the activity recognition, which overperformed the classical machine learning methods. [8] proposed a new structure which combined CNN and RNN (DeepConvLSTM) to realize a higher performance compared with baseline CNN for HAR. In 2017, [9] presented three DRNN (Deep-RNNs) architectures, which outperformed the other state-of-the-art methods including

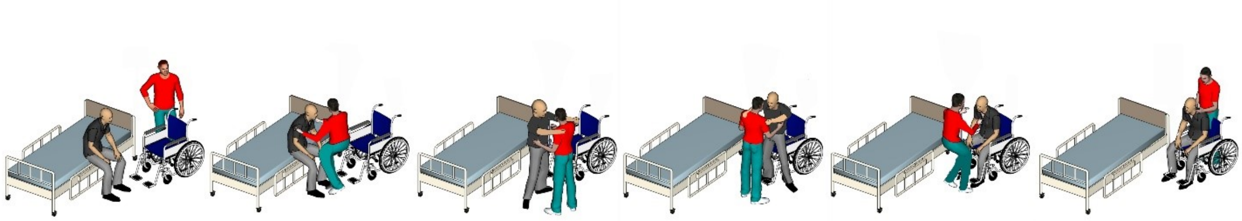

Fig. 1. The process of patient transfer

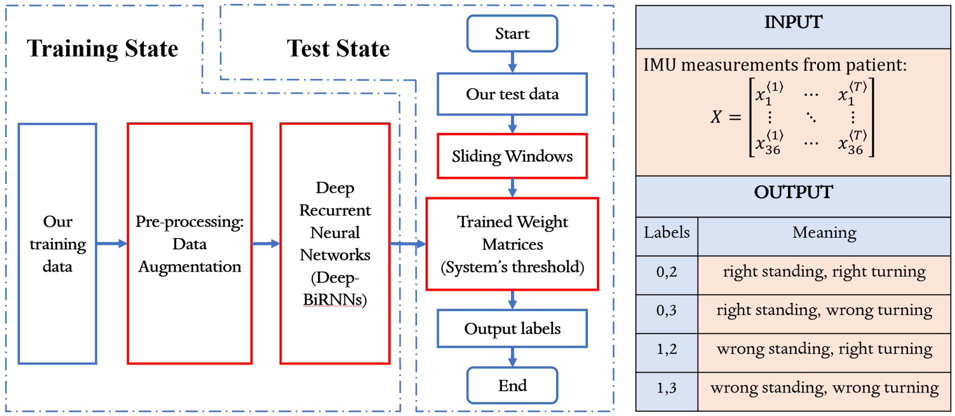

Fig. 2. Structure of Proposed System

DeepConvLSTM. Then, in 2018, [10] borrowed the idea from Microsoft's ResNet (Residual Network) and applied it on RNN structure (ResBidirLSTM), the performance of which was better than baseline CNN and RNN.

However, the datasets used by above related works are all about activities of daily living, including the actions class like running, walking, jumping, lying, sitting, upstairs, downstairs and so on. Professional actions like nursing care has not been explored yet. In order to evaluate the performance of the nurse student, it's necessary to realize the recognition for these professional actions, which involve very similar action class like a correct and an incorrect way of a same action. No one has ever explored the performance of machine learning methods or deep learning methods for such special HAR task, which is more difficult to realize the recognition compared with human activities of daily living. As for continuous professional nursing actions, the effectiveness of conventinal sliding windows method is unknown.

III. Proposed System

A. Precondition

The process of patient transfer will be our system's target actions Fig.1, since patient transfer is one of heaviest and most difficult nursing activities, which needs nurse to use proper body mechanics with appropriate timing of the transfer [11]. Furthermore, patient transfer is performed frequently by nurses in hospital and nursing homes [12].

This study will mainly focus on two steps, standing and turning of patient transfer process. Because these steps are involving a large range of motion, which is easy to get hurt for both nurse and patient and is hard to learn for nursing students. And it's convenient to expand from these two consecutive steps to the whole process.

There will be one right and one typical wrong method for each action as Tab.I. These correct and incorrect methods corresponding to the critical steps were discussed and chosen by professional nurse teachers. These incorrect methods are most common mistakes made by nurse students.

TABLE I
Taget action type

	Right	Wrong
Standing	Bending waist	Without Bending waist
Turning	Using left foot	Using right foot

B. System Architecture

The proposed system includes training state and test state in Fig.2. Firstly, we will preprocess the training data and feed them into our Deep-RNNs model in the training state. After training, the model could realize the recognition for separate action. As for continuous actions, sliding windows will be applied for segmenting the data before testing.

C. Deep Recurrent Neural Networks

Recurrent neural networks (RNNs) will be used to build our system's classification, which is the sequence model in Deep Learning, specialized for dealing with time-series signal. Compared with CNN, RNN with LSTM cell has the ability to capture a long temporal correlation between data samples, while the size of CNN kernels restricts the captured range of dependencies between data samples. The mathematical process of standard RNN with LSTM cell is as below:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\widetilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * tanh(C_t) \tag{6}$$

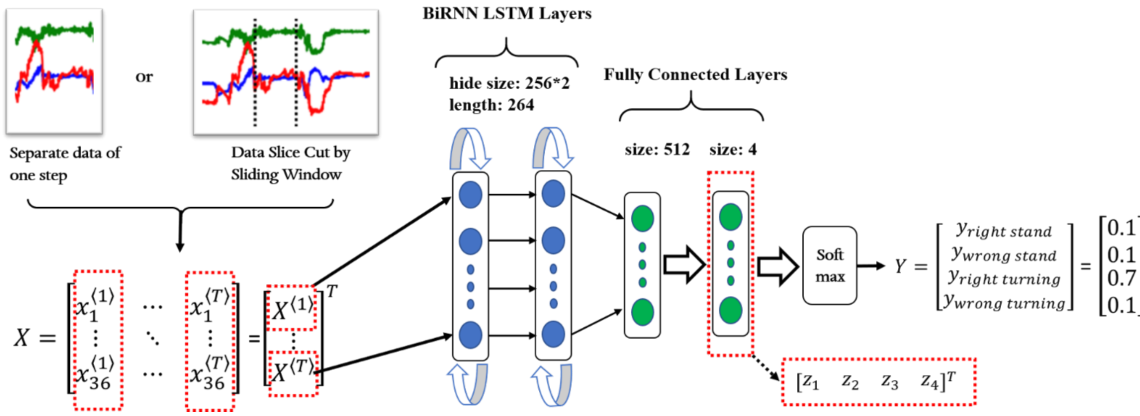

Fig. 3. Implementation of Deep-RNNs in Our System

Stack multiple RNNs layers of LSTM cells with bidirection can build a Deep-RNNs model, which could learn more complex mapping function from input to output for time-series signal [13] [14]. Due to the strong ability of Deep-RNNs for extracting correlation among data samples, it could directly deal with raw data without designing and extracting hand-crafted features, and realize the recognition of action regardless of the influence of individual difference [15].

The real implementation of Deep-RNNs is as Fig.3. There are two bidirectional LSTM RNN layers with 512 hidden size. Then the output is connected by fully connect layers with 4 output nodes, the number of action classes. Finally, through SoftMax function,

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad for\ i = 1, ..., K \tag{7}$$

the system could obtain the probability of each action class.

D. Sliding Windows

For real application, the input data will be continuous actions. Thus, we need to use sliding windows as Fig.4 to get data fragments as the input of our network. According to the results for each data fragment, markers will be attached to corresponding time point. By dealing with the markers, the evaluation system can get final evaluation results.

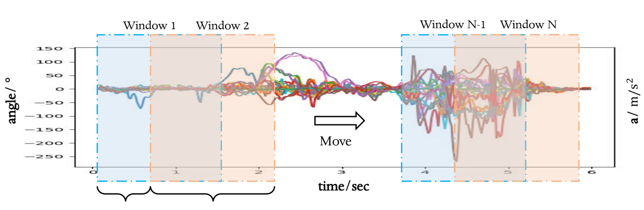

Fig. 4. Sliding windows

E. Data Augmentation

Instead of using the above system to directly deal with the data fragment from sliding windows, data augmentation (cropping) was used to reduce the variance of the classifier.

IV. Experiment

A. Purpose

We did a experiment by employing students with patient transfer skill first. The purpose of this experiment is as follows:

- Build a prototype evaluation system first and verify if the proposed model can work or not;
- Get a general feeling about the demand for data and find some problems that need to be pay attention to when we cooperate with real nursing staffs.

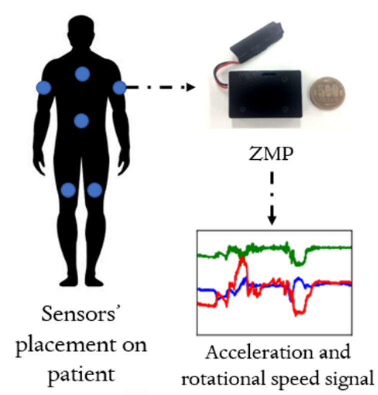

Fig. 5. Placement of Sensors

B. Experiment Setting

Totally 6 IMU sensors (ZMP) were installed on different parts of patient's body, including arm, thigh, chest and abdomen as Fig.5. Each sensor has 6 signal channel outputs:

- 3 acceleration channels

- 3 rotational speed chaneels

Thus, there are totally 36 $6 \cdot (3+3) = 36$ channels collected from patient's body. The work frequency of the sensor is at 50Hz.

C. Experiment Content

TABLE II
Experiment Content

Content	Nurse1	Nurse2	All
Right Standing	20	20	40
Wrong Standing	20	20	40
Right Turning	20	20	40
Wrong Turning	20	20	40
Right Standing and Right Turning	4	4	8
Right Standing and Wrong Turning	4	4	8
Wrong Standing and Right Turning	4	4	8
Wrong Standing and Wrong Turning	4	4	8

In this experiment, every nurse (experienced student) will do each kind of action class 20 times (Standing and Turning), including a right way and a typical wrong way. And each nurse would do every possible continuous actions 4 times. The Experiment content is in Tab.II. The real situation of this experiment is shown in Fig.6.

Fig. 6. Real Situation of Experiment

V. Results

We used PyTorch platform to build our model. Adam optimizer and cross entropy loss function (8) were used,

$$Loss(x, class) = -log\left(\frac{e^{x[class]}}{\sum_j e^{x[j]}}\right) \quad (8)$$

A. Results of Separate Action

As for separate action, the ratios for train, valid and test dataset are as follows:

- training set: 60%
- valid set: 20%
- test set: 20%

During the training process, the accuracy for valid set can finally reach 93.75% Fig.7, while the accuracy for test set is 87.50%. The concrete prediction results for each separate action class can see the confusion matrix in Fig.8.

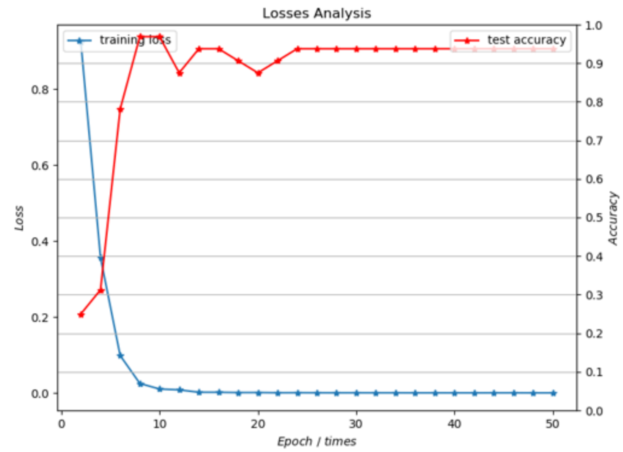

Fig. 7. Accuracy and Loss Analysis

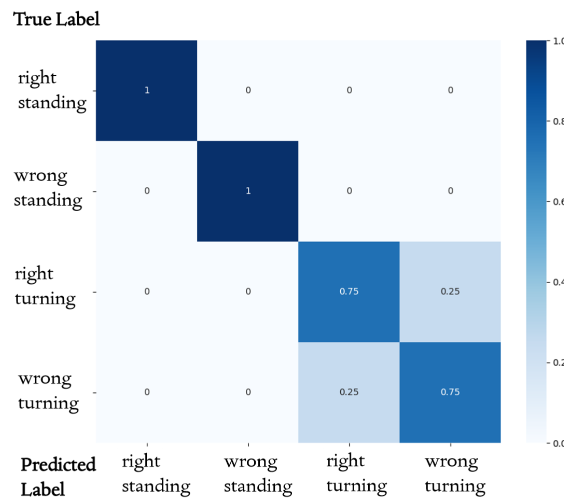

Fig. 8. Confusion Matrix for Separate Action

B. Results of Continuous Action

As for continuous actions, when making the evaluation, the action sequence was assumed as correct. Then, the best model can reach 89% accuracy by applying sliding windows. However, the model after training for continuous actions test is unstable. Usually, the judgment for standing (0 or 1) is right, but for turning (2 or 3), the result is bad. On average, the system can only reach 75% accuracy. The concrete prediction results for each continuous actions class can see confusion matrix in Fig.9.

VI. Discussion

From the results, we can see that Deep-RNNs can make the judgment for the professional nursing actions without designing and calculating hand-crafted features. Furthermore, it can be robust to the influence of individual differences of nurses.

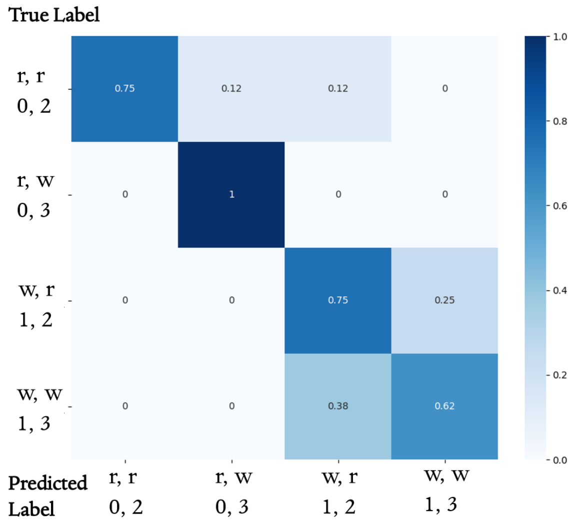

Fig. 9. Confusion Matrix for Continuous Actions

The possible reason why proposed system cannot handle continuous actions well was discussed. Sliding windows are implemented in all related works in HAR field for dealing with continuous data and can get relative good results. However, according to action classes in public databases, many activities of daily living belong to “repetitive structure”, like running, walking, upstairs and downstairs. Because of their repetitive attribution, they are robust to the size and location of sliding windows. However, because of professional actions’ special attribution, they may be very sensitive to the size and location of the windows. Especially for action step like turning, the difference between right and wrong method of turning is so subtle that the system can easily make a wrong judgment if windows were not in a good location and with a good size.

VII. Conclusion

In conclusion, we verify the performance of Deep-RNNs model for recognizing professional nursing actions. The results show that Deep-RNNs can effectively make the judgment for very similar separate action and is robust enough to the influence of individual difference of different nurses. However, sliding windows method cannot work well for such continuous professional actions. A new method of processing continuous data for such professional actions are required.

References

[1] Grady, Ptricia A. ”Advancing the health of our aging population: A lead role for nursing science.” Nursing outlook59.4 (2011): 207.

[2] Allan, Janet D., and Jillian Aldebron. ”A systematic assessment of strategies to address the nursing faculty shortage, US.” Nursing outlook 56.6 (2008): 286-297.

[3] Nardi, Deena A., and Charlene C. Gyurko. ”The global nursing faculty shortage: Status and solutions for change.” Journal of Nursing Scholarship 45.3 (2013): 317-326.

[4] Yang, Jianbo, et al. ”Deep convolutional neural networks on multichannel time series for human activity recognition.” Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015.

[5] Jiang, Wenchao, and Zhaozheng Yin. ”Human activity recognition using wearable sensors by deep convolutional neural networks.” Proceedings of the 23rd ACM international conference on Multimedia. Acm, 2015.

[6] Gjoreski, Hristijan, et al. ”Comparing deep and classical machine learning methods for human activity recognition using wrist accelerometer.” Proceedings of the IJCAI 2016 Workshop on Deep Learning for Artificial Intelligence, New York, NY, USA. Vol. 10. 2016.

[7] Tao, Dapeng, Yonggang Wen, and Richang Hong. ”Multicolumn bidirectional long short-term memory for mobile devices-based human activity recognition.” IEEE Internet of Things Journal 3.6 (2016): 1124-1134.

[8] Ordóñez, Francisco, and Daniel Roggen. ”Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition.” Sensors 16.1 (2016): 115.

[9] Murad, Abdulmajid, and JaeYoung Pyun. ”Deep recurrent neural networks for human activity recognition.” Sensors17.11 (2017): 2556.

[10] Zhao, Yu, et al. ”Deep residual bidir-LSTM for human activity recognition using wearable sensors.” Mathematical Problems in Engineering 2018 (2018).

[11] Fash, Bernice, and Frances Powell. ”Body mechanics in nursing arts.” The American Journal of Nursing (1941): 190-195.

[12] Garg, A., B. D. Owen, and B. Carlson. ”An ergonomic evaluation of nursing assistants’ job in a nursing home.” Ergonomics 35.9 (1992): 979-995.

[13] Hochreiter, Sepp, and Jürgen Schmidhuber. ”Long short-term memory.” Neural computation 9.8 (1997): 1735-1780.

[14] Olah, Christopher. ”Understanding lstm networks.” (2015).

[15] Wang, Jindong, et al. ”Deep learning for sensor-based activity recognition: A survey.” Pattern Recognition Letters 119 (2019): 3-11.