IEEE Sensors Letters VOL. 7, NO. 10, OCTOBER 2023

IEEE Sensors Council

Sensor systems

Human Activity Recognition Using 2-D LiDAR and Deep Learning Technology

Qiao-Yuan Yao, Po-Lin Chen, and Tzung-Shi Chen*

Department of Computer Science and Information Engineering, National University of Tainan, Tainan 70005, Taiwan
**Member, IEEE*

Manuscript received 25 July 2023; accepted 13 September 2023. Date of publication 19 September 2023; date of current version 27 September 2023.

Abstract—In this letter, a home care assistance system is proposed for elder human activity recognition that combines 2-D LiDAR and deep learning technology. First, the 2-D LiDAR is used to scan the room's interior data, and cluster algorithms are used to identify high-density areas that may be objects. The resulting data are then classified into human and nonhuman clusters. The coordinates of the human clusters are sequentially recorded to generate trajectory graphs. These trajectory graphs possess both spatial and temporal attributes and are processed using spatial–temporal graph convolutional networks to achieve accurate classification of human activities. Furthermore, the system is designed to detect abnormal trajectories that may indicate a fall and can send out a warning signal to notify the caregiver for emergency assistance. With these advanced features, the home care assistance system can improve the safety and well-being of elderly individuals living alone or with limited assistance.

Index Terms—Sensor systems, 2-D LiDAR, deep learning, home care, human activity recognition (HAR), machine learning.

I. INTRODUCTION

The aging of the population is an important demographic trend that is affecting countries around the world. As fertility rates decline and life expectancy increases, the proportion of older people in the population is increasing. According to the United Nations' World Population Prospects 2022 [8], the proportion of people aged 65 and over is expected to increase from 10% in 2022 to 16% by the middle of this century. By 2050, the number of people aged 65 and over is expected to be more than twice the number of children under 5 worldwide. This shift in age distribution presents challenges for countries, particularly in the area of public planning and the provision of long-term care systems for the elderly. Developing countries face even greater challenges in adapting to the aging of their populations, as they are often struggling to provide adequate care for older people. This is where home care assistance systems can play a critical role in providing support and assistance to older adults in need. By implementing home care assistance systems, developing countries can help to ensure that elderly people receive the care and support they need to maintain their independence and quality of life [12].

In recent years, there has been an increase in the development of home care assistance systems that help caregivers monitor the daily activities of care recipients and reduce their burden. Human activity recognition (HAR) is a critical technology in these systems, as it allows appropriate sensors to be selected based on the environment and assists in monitoring one or more individuals in different locations [1], [7]. Different sensors are subject to limitations, such as light, detection distance, angle, weather, etc. While cameras provide high precision, they require sufficient lighting to function effectively and may raise concerns about privacy breaches [6], [9]. In contrast, LiDAR sensors offer the advantage of preventing privacy breaches while maintaining accuracy in varying lighting conditions and have good real-time performance. Therefore, in this letter, we have selected LiDAR as the sensor for the home care system.

Deep learning has gained popularity in recent years due to its remarkable ability to process complex data, making it a popular research topic across many fields. In the case of LiDAR, it is used to collect sequential data. While recurrent neural networks (RNNs) were traditionally the preferred model for sequential analysis, the advent of long short-term memory (LSTM), a variant of RNN, has mitigated the problem of gradient exploration or vanishing and has since become a popular choice [3], [9].

In this letter, 2-D LiDAR is used to collect data in the form of point cloud. We first segment objects using density-based spatial clustering of applications with noise (DBSCAN) [4] and then use random forest (RF) [2] for classification as preliminary data preprocessing. Finally, we apply the spatial–temporal graph convolutional network (ST-GCN) model proposed by Yan et al. [11] to fully utilize spatial and temporal features for HAR. Our contributions are as follows.

1) We combine the characteristics of LiDAR point cloud data and use ST-GCN as the classification model, achieving good accuracy.
2) Our proposed system can predict the abnormal trajectory of human behavior caused by falls and assist in issuing warnings while ensuring the privacy and security of the care recipient for home care.

The rest of the letter is organized as follows. Section II covers the preliminaries of this letter. Section III presents the proposed method and experimental results. Section IV describes the design and implementation of the system. Finally, Section V concludes this letter.

II. PRELIMINARIES

In order to distinguish objects, the data collected by LiDAR must first be clustered before further analysis [5]. DBSCAN is used for grouping point clouds in a feature space based on density [4], [13]. To implement DBSCAN, two parameters need to be defined, which are as follows.

1) ε (*eps*): The circular area drawn with this parameter as the radius is called ε-neighborhood.

Corresponding author: Tzung-Shi Chen (e-mail: chents@mail.nutn.edu.tw).
Associate Editor: S. Dhanekar.
Digital Object Identifier 10.1109/LSENS.2023.3316882

Table 1. Eight Features Extracted for RF Clustering

No.	Geometric features	No.	Geometric features
1	Length	2	Width
3	Standard deviation of the X coordinate	4	Standard deviation of the Y coordinate
5	Variation of the X coordinate	6	Variation of the Y coordinate
7	The average distance from each point to the center	8	Number of points

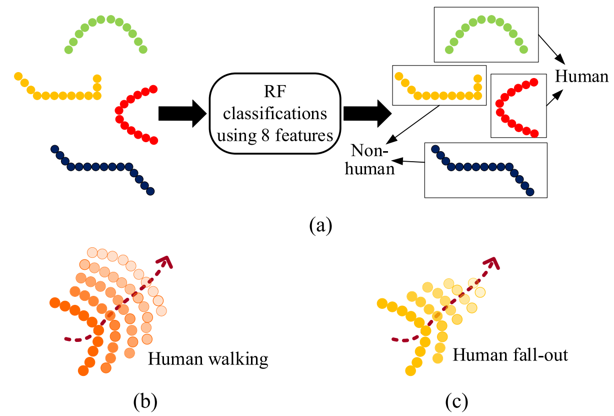


Fig. 1. Human moving trajectories. (a) Human identified using RF. (b) Normal trajectory. (c) Abnormal trajectory.

2) *minPts*: This is the minimum number of points required to define an area as high density.

After applying DBSCAN to the data points, they are grouped into various clusters, representing different objects, such as walls, cabinets, and humans in the room. In order to track human trajectories, it is necessary to identify the cluster that represents humans [5], [10]. To achieve this, we extracted eight features, as depicted in Table 1, and utilized RF to classify the objects as either human or nonhuman by comparing the geometric features of each cluster.

RF is a classifier consisting of multiple decision trees that use bagging and random feature sampling. To classify objects as human or nonhuman, RF outputs the class selected by the majority of decision trees [5]. This classifier is effective in eliminating outliers and making predictions. In this letter, we use the clusters that have been labeled as human or nonhuman as the training set to train the RF classifier to output the binary results we need, as shown in Fig. 1(a). Hereby, out of the 6000 samples used for training, 3000 were human and 3000 were nonhuman.

To minimize the chances of false positives, the RF classifier may mistakenly identify other objects in the room as humans, resulting in multiple individuals being identified at the same time. A tracker is used to calculate the distance between all objects classified as humans and their previous location. The individual with the closest distance is selected as the current human. In cases where no human was detected in the previous period, the first human object identified is chosen for tracking. The moving trajectories with eight consecutive point clouds (time frames) for humans moving from the bottom left corner to the top right corner are shown as in Fig. 1, while 2-D LiDAR was placed on the right-hand side. In Fig. 1(b), a normal trajectory is shown. Fig. 1(c) shows the trajectory for the human fall-out as time goes on.

Yan et al. [11] proposed ST-GCN for active human recognition using skeletal points in a continuous temporal sequence. Here, ST-GCN is utilized for recognizing human activities in LiDAR point clouds. Each recognized human point cloud in time t is an undirected

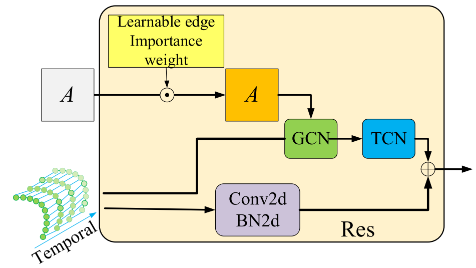


Fig. 2. ST-GCN model structure we used.

graph $G_H^t = (V_H^t,\ E_H^t)$, where $V_H^t = \{v_1^t,\ v_2^t,\ \ldots,\ v_{h_t}^t\}$, with consecutive points $v_1,\ v_2,\ \ldots,\ v_{h_t}$, via 2-D LiDAR in clockwise scanning, for $1 \le t \le T$, and T is the total time steps when moving. $E_H^t = \{\ e_i^t = (v_i^t,\ v_{i+1}^t) | 1 \le i < h_t\}$. The spatial–temporal graph for humans moving with N_f, a number of time frames, from t to $(t + N_f - 1)$, of cloud points, is defined as an undirected graph $G_{HT}^t = (V_{HT}^t,\ E_{HT}^t)$, where $V_{HT}^t = \bigcup_{i=t}^{t+N_f-1} V_H^i$, and $E_{HT}^t = \bigcup_{i=t}^{t+N_f-1} E_H^i \cup$

$$\bigcup_{i=t}^{t+N_f-2}\left(\begin{array}{c}\left\{e_j^i=\left(v^i_{\left\lceil j\times\frac{h_i}{h_{i+1}}\right\rceil},\ v_j^{i+1}\right)\middle|\,1\le j\le h_{i+1},\ \text{ if } h_i\le h_{i+1}\right\}\\ \text{or}\\ \left\{e_j^i=\left(v_j^i,\ v^{i+1}_{\left\lceil j\times\frac{h_i}{h_{i+1}}\right\rceil}\right)\middle|\,1\le j\le h_i,\ \text{ if } h_i> h_{i+1}\right\}\end{array}\right).$$

The model begins by creating a weight matrix A based on G_{HT}^t for edge importance and using it to weight the adjacency matrix. The weighted matrix and trajectory map are then sent to the graph convolutional network (GCN) for calculation, and a residual structure is used to obtain Res, which is combined with the output of the temporal convolutional network (TCN) to aggregate time-dimensional features. The used model is shown in Fig. 2. The 2-D convolution layer (also known as Conv2D) and 2-D Batch Normalization layer (also known as BN2D) are components within the ST-GCN model.

III. METHOD AND EXPERIMENTS

The proposed method was carried out in three steps, as shown in Algorithm 1. To ensure a more accurate dataset, five different paths, five different stopping positions, and a falling-down scenario were collected. The abnormal activities were when a participant fell down while walking on the path because it was unclear which destination to choose. In total, there were 11 types of activities, each lasting 7 s from beginning to end. The execution button was pressed, and the walking behavior started 2 s later and continued for 5 s [3].

In the first step of this research, clustering was performed using DBSCAN to group the points into different objects. The various groups represented different furniture. In the second step, the grouped objects were classified as human or nonhuman using the RF algorithm. A tracker was employed to continuously track the object identified as human. If there were more than two people in the room, the tracker found the shortest distance from the previous person to track. Finally, ST-GCN was used to classify the activities corresponding to the trajectories. The effectiveness of the model was evaluated, and its execution time and accuracy were compared with those of TCN and LSTM.

In Algorithm 1, a 2-D LiDAR point cloud in polar coordinates (r, θ) is collected. The data are first converted from polar coordinates to Cartesian coordinates and stored in an array. The LiDAR data collection continues until *maxPoints* (set to 720) points are collected,

Algorithm 1: HAR for Home Care.

```
Input: 2D LiDAR point cloud in polar coordinates (r, θ) at each time t.
Output: The class of human activity corresponding to the trajectory.
While True do
    scanningMap←[]; groupedPoints←[];
    numPointsCollected←0
    For i from 1 to numframe
        While numPointsCollected < maxPoints do
            polarCoords ← collectLiDARData()
            cartCoords ← polarToCartesian(polarCoords)
            scanningMap.append(cartCoords)
            numPointsCollected ← numPointsCollected + 1
        End while
        clusteringResults ← DBSCAN(scanningMap)
        groups ← grouping(scanningMap, clusteringResults)
        For all group ∈ groups do
            features.append(calculateFeatures(group))
        End for
        trajectories.append(RF(features))
    End for
    predict ← ST-GCN_MODEL(trajectories)
End while
```

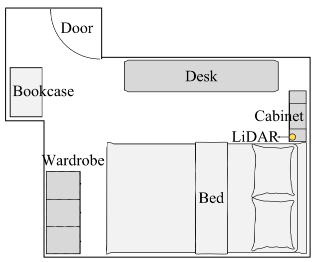


Fig. 3. Layout of the test site.

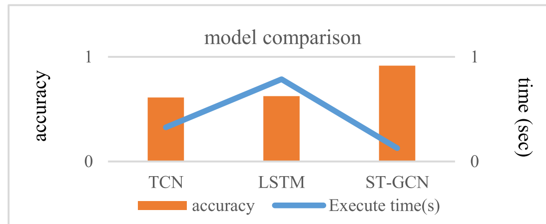
which we consider to be sufficient to capture the shape and trajectory of a human while maintaining the real-time performance of the system. Once *maxPoints* points are collected, the scanning outline map is clustered using the DBSCAN algorithm, resulting in an array of size 720 that records the cluster label of each point in the scanning outline map. The points are then grouped using the grouping function to obtain different groups, which could represent furniture or human subjects. Next, features are calculated for each group, and an RF classifier is used to classify which group represents a human subject. This process is repeated *numframe* (set to 15) times, and the resulting "human" datasets are combined to form trajectories. The trajectories are then input into an ST-GCN model, which has six classes, to classify human activity. Overall, this algorithm provides a method to identify and track human activity using LiDAR data. The use of DBSCAN for clustering and RF for classification allows for efficient and accurate identification of human subjects within the scanning outline map. The ST-GCN model further provides a means of classifying the human activity based on the trajectory.

The layout of the test site is shown in Fig. 3, and it consists of a bedroom equipped with a bed, a desk, a bookcase, and a wardrobe. A 2-D LiDAR is placed on a cabinet approximately 120 cm tall to scan the dynamic situation in the room. It is important to avoid furniture overlapping as the furniture arrangement and aisle can affect the scanning situation. The LiDAR used in the study is the Slamtec RPLiDAR A1, which has a low cost and a scanning range of about 12 m in radius. It is embedded in a logic IO driver (3.3 V) motor controller, which can be configured by adjusting the motor speed to set the scanning frequency. To begin scanning, the RPLiDAR A1 should be connected to the computer with a MicroUSB cable.

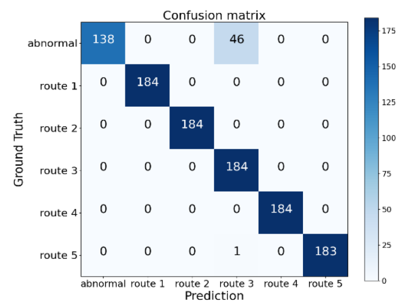
Here, there are seven different trajectories for human activity in a test site, as depicted in Table 2. The accurate recognition of human activity is crucial for many systems, and data augmentation can play a key role in increasing the amount of available data by generating slightly modified copies of existing data. This study utilizes three different models for testing and comparison purposes, with 80% of the machine learning data (10 800 records) allocated to the training set, 10% to the validation set, and 10% to the test set. Table 2 depicts the running time and accuracy for different models.

Table 2. Seven Different Trajectories for Human Activity

Trajectory	Source	Destination	Human activity
Abnormal	Anywhere	In room	Fall-Out
Route 1	In front of the bed	Wardrobe	Change clothes
Route 2	The left side of the bed	Desk	Reading
Route 3	Desk	Bookcase	Take book
Route 4	In front of the bed	Door	Get up
Route 5	Desk	Wardrobe	Ready to take a shower



(a)



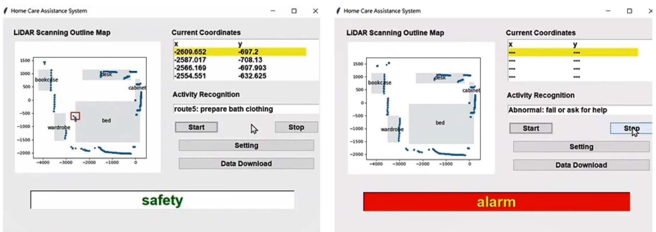
(b)

Fig. 4. Model comparison and confusion matrix for classification. (a) Model comparison for accuracy and execution time. (b) Confusion matrix for classification in ST-GCN.

Based on the experimental results, it was found that ST-GCN outperformed TCN and LSTM in terms of accuracy and execution time as in Fig. 4. This is because TCN captures temporal features while GCN captures spatial features. The model comparison is shown in Fig. 4(a).

According to the results of the model comparison, ST-GCN was found to be the best model for HAR. The classification results for the model were represented by the confusion matrix shown in Fig. 4(b). The rows in the matrix represented the actual labels while the columns represented the predicted labels. The numbers in the boxes represented the results of the ST-GCN model. One data point for route 1 was misclassified as abnormal, indicating that the model had a high level of accuracy. In Fig. 4(b), abnormal activities were mistakenly classified as route 3 occurrences 46 times. The main reason is that when a person moves away from the position of the 2-D LiDAR, this behavior resembles falling. The fall-out behavior resulted in a higher rate of misclassification as route 3.

To evaluate the performance of ST-GCN in multiclass classification tasks, we employed several performance metrics. Our experimental results showed that the model achieved an accuracy of 96%, indicating a high overall correctness in sample classification. Furthermore, the weighted average precision calculated for all classes resulted in a



(a) (b)

Fig. 5. Snapshots of the proposed home care assistance system. (a) Person in safety. (b) Alarm occurred while a person fall-out.

value of 97%. The weighted average recall obtained a value of 96%. Finally, the weighted average F1 score yielded a value of 96%. These metrics reflect the model's ability to accurately classify samples while considering class distribution. The experimental results demonstrate the outstanding performance of our model in multiclass classification tasks.

IV. SYSTEM IMPLEMENTATION

In this letter, we designed and implemented an assistance system with a user-friendly UI interface to make it more accessible to new users and seniors. The interface displays all operable functions in a graphical format, as shown in Fig. 5(a), and is designed to ensure the safety of the care recipient.

Users can initiate the scanning process by clicking the "Start" button, which updates the LiDAR scanning outline map in real time on the left side of the interface. The system employs DBSCAN and RF algorithms to process the collected data and classify it into two groups: 1) human and 2) nonhuman. The current position of the elderly is displayed in the "Current Coordinates" section and simultaneously marked on the LiDAR scanning outline map. The system calculates the care recipient's current path and corresponding behavior every 15 time frames and records the information in the "Activity Recognition" section. The status bar at the bottom of the interface indicates whether the current status is "safe" or "warning." If the status is "warning," the system emits an alert sound to notify caregivers to provide assistance. Users can stop the scanning process and updates by clicking the "Stop" button.

When a care recipient falls, the system will display an immediate alert notification, as shown in Fig. 5(b). Caregivers can view the notification on their own devices or through the central monitoring system. The notification will include the care recipient's current location, a description of the situation, and the appropriate action to take. The system may also automatically contact emergency services or designated contacts if the situation warrants it. The interface may display additional information, such as the care recipient's medical history or emergency contact information, to help caregivers respond appropriately.

The proposed system can be deployed not only at home but also installed in embedded devices in Internet-of-Things environments. The system can be widely deployed in multiple rooms and can even be used in wet and slippery places or swimming pools. Through system integration and development, it can achieve comprehensive privacy health care.

V. CONCLUSION

This letter described an assistance system designed to improve convenience and provide caregivers with more flexibility while ensuring the home care recipient's privacy and security. The ST-GCN model was used and demonstrated the fastest execution time compared to other models, and its accuracy is significantly higher. The system displays the current position of the elderly person on the scanning outline map and records their trajectory and corresponding behavior in real time. Finally, whenever a warning status is detected, the system alerts caregivers with an alarm.

ACKNOWLEDGMENT

This work was supported by the National Science and Technology Council under Grant NSTC-111-2813-C-024-012-E and Grant NSTC-110-2221-E-024-003-MY3, Taiwan.

REFERENCES

[1] M. Bouazizi, C. Ye, and T. Ohtsuk, "2-D LIDAR-based approach for activity identification and fall detection," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 10872–10890, Jul. 2022.

[2] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.

[3] T.-C. Chiu, T.-S. Chen, and J.-M. Lin, "Human gait recognition using LiDAR and deep learning technologies," in *Proc. IEEE Int. Conf. Consum. Electron. - Taiwan*, 2022, pp. 43–44.

[4] D. Deng, "DBSCAN clustering algorithm based on density," in *Proc. 7th Int. Forum Elect. Eng. Automat.*, 2020, pp. 949–953.

[5] F. Luo, S. Poslad, and E. Bodanese, "Temporal convolutional networks for multiperson activity recognition using a 2-D LiDAR," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7432–7442, Aug. 2020.

[6] B. Parsa and A. G. Banerjee, "A multi-task learning approach for human activity segmentation and ergonomics risk assessment," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2021, pp. 2352–2362.

[7] J. Roche, V. De-Silva, J. Hook, M. Moencks, and A. Kondoz, "A multimodal data processing system for LiDAR-based human activity recognition," *IEEE Trans. Cybern.*, vol. 52, no. 10, pp. 10027–10040, Oct. 2022.

[8] "World Population Prospects 2022: Summary of results," Tech. Rep. UN DESA/POP/2022/TR/NO. 3, Population Division, United Nations Dept. Econ. Social Affairs, New York, NY, USA, 2022.

[9] Q. Wang, K. Zhang, and M. A. Asghar, "Skeleton-based ST-GCN for human action recognition with extended skeleton graph and partitioning strategy," *IEEE Access*, vol. 10, pp. 41403–41410, 2022.

[10] W. Wang et al., "PointLoc: Deep pose regressor for LiDAR point cloud localization," *IEEE Sensors J.*, vol. 22, no. 1, pp. 959–968, Jan. 2022.

[11] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Proc. Assoc. Advance. Artif. Intell. Conf. Artif. Intell.*, 2018, vol. 32.

[12] Y. Yang, H. Wang, R. Jiang, X. Guo, J. Cheng, and Y. Chen, "A review of IoT-enabled mobile healthcare: Technologies, challenges, and future trends," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9478–9502, Jun. 2022.

[13] Z. Yu et al., "A radar-based human activity recognition using a novel 3-D point cloud classifier," *IEEE Sensors J.*, vol. 22, no. 19, pp. 18218–18227, Oct. 2022.