Leg Detection for Socially Assistive Robots: Differentiating Multiple Targets with 2D LiDAR

Hanchen Yao[1,2,3], Jianwei Peng[1,2,3], Zhelin Liao[2,3], Ran Zhao[4], and Houde Dai[1,2,3](✉)

[1] Fujian College, University of Chinese Academy of Sciences, Jinjiang 362216, China
{yaohanchen21,pengjianwei20}@mails.ucas.ac.cn
[2] Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou 350002, China
dhd@fjirsm.ac.cn
[3] Quanzhou Institute of Equipment Manufacturing, Haixi Institutes, Chinese Academy of Sciences, Jinjiang 362216, China
3211239023@fafu.edu.cn
[4] Zhongyuan-Petersburg Aviation College, Zhongyuan University of Technology, Zhengzhou 451191, China
zhaoran@zut.edu.cn

Abstract. While socially assistive robots working in environments with a lot of people walking and obstructions, LiDAR-based detectors may have trouble locating the target person. The lack of identifying information is a drawback of the 2D range data obtained by a LiDAR sensor. Consequently, when applying the traditional technique of clustering 2D laser dots with geometric properties, modern LiDAR-based leg detectors typically fail. To recognize and identify the target individual in real time, an improved leg detector based on density-weighted support vector data description (DW-SVDD) is presented. The suggested DW-SVDD leg detector in this study incorporates density weight, width, and girth features into the support vector data description. Socially assistive robots that follow humans may quickly and accurately identify items from vast quantities of 2D laser point data with this detector. To evaluate detection accuracy, the proposed leg detector is tested on a mobile robot both indoors and outdoors. Field experiment results show that the proposed DW-SVDD leg detector enables socially assistive robots to detect partial occlusions and similar obstacles effectively. Additionally, the results indicate that the proposed leg detector achieves an MOTA of 34.96% and MOTP of 39.17%, demonstrating its efficacy in practical applications.

Keywords: Leg detection · Socially assistive robot · Multiple target detection

This work was supported in part by the Central Government Guides Local Special Funds for Science and Technology Development under Grant 2020L3028 and 2021L3047.

F. Sun et al. (Eds.): ICCSIP 2023, CCIS 1918, pp. 87–103, 2024.
https://doi.org/10.1007/978-981-99-8018-5_7

1 Introduction

Is there a solution to specifically identify the target in a multi-person scenario without the use of wearable sensors? Human-robot interaction (HRI) requires socially assistive robots to detect, identify, and track the behavior of the target person [1]. There is an urgent requirement for socially assistive robots [2–4] to serve a congested workspace due to the acceleration of aging process and the rise in labor costs (e.g., hospital, supermarket, restaurant, or warehouse). In multi-person and cluttered scenarios, identifying interactive targets without the use of wearable sensors remains an unresolved challenge.

Camera sensors have gained popularity in the robotics market. However, they lose the ability to recognize pedestrians at night. Compared to RGB-D cameras require advanced computational power and violate personal privacy [5], Light Detection And Ranging (LiDAR) is a high-accuracy range sensor that provides 2-dimensional (2D) range information at high rates for dynamic environments. The usage of RGB-D cameras is constrained by the COVID-19 pandemic since masks hide a significant amount of RGB information about the face. Additionally, RGB-D cameras are rarely useful in environments with inadequate or excessive lighting when used outside.

A LiDAR-based socially assistive robot, which can serve as a logistical help, doesn't need active markers or optical reflection markers. For instance, the human-following robot EffiBOT (Effidence Inc., Romagnat, France) uses LiDAR sensors [7]. EffiBOT can follow a picking operator due to its "Follow-me" functionality. The efficiency of logistics in factories and warehouses is increased because picking operators no longer need to manually push or pull carts.

The majority of researches concentrated on leg detection with LiDAR because of the height restrictions of mobile robots. Compared to waist detection [8] and shoulder detection [9], leg detection does not require adjusting the measurement range based on the person's height. Leg detection at a lower height can simultaneously detect human legs and obstacles, while waist detection and shoulder detection at a higher height are considered to be prone to overlooking potential obstacles. Generally speaking, investigations [8–12] aim to build a classifier based on the leg's geometry constraints. The circle fitting [9] and bounding box [8] are traditional methods for obtaining leg characteristics. However, neither of the two approaches is sufficient to keep the leg detecting stable. Arras *et al.* [10] created a ground-breaking work in the AdaBoost approach to develop a robust classifier, which collected 14 characteristics to learn the shape of legs, to address the issue brought on by the sparsity of the LiDAR data. The drawback is that the situation with the occluded legs was not taken into account. Li *et al.* [11] added extra characteristics to an AdaBoost-based detector to increase the detection accuracy when legs are partially occluded. Additionally, Beyer *et al.* [12] used the leg features to create a deep learning (DL) detector. This DL method's lengthy run-time makes it impossible to employ online with a robot that follows people. In conclusion, these methods are prone to false positives brought on by ambient noises, such as chair and table legs. Due to the poor information quality

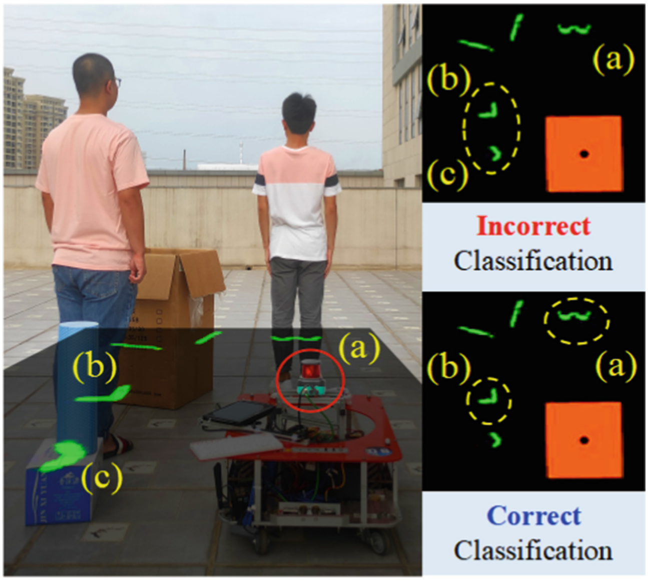


Fig. 1. In multi-person environments, leg detection may experience recognition errors due to insufficient features. The test dataset is divided into incorrect classification and correct classification. (a) Laser data where both legs are connected is difficult to recognize as a single person. (b) Side-standing leg data is obstructed by the other leg. (c) Obstacles resembling cylinders are often mistakenly identified as human legs.

of 2D range data, it is difficult for these conventional classifiers to categorize the target person's leg in the crowd.

Recently, to solve the classification issue in 2D range data, some researchers used the support vector data description (SVDD) as a one-class classification (OCC) [13]. Chung *et al.* [14] used an SVDD-based classifier to choose leg clusters without presuming that human legs have any particular geometric shapes. As a result, even though more complicated leg forms were supported for searching, the accuracy of leg detection decreased. Cha *et al.* [15] addressed this problem by modifying the SVDD-based classifier to learn the classification boundary in 3D feature space using a 2D LiDAR. Jung *et al.* [16] used an SVDD algorithm to detect the shoulder using a marathoner service robot in order to assess the outside performance. Despite the great performance of these SVDD algorithms for detecting the target person's legs, it was challenging to distinguish a specific target from other objects and occlusions. The SVDD-based leg detector could identify hallways and human legs for side-by-side following in our prior investigation [17] (Fig. 1).

This paper suggests an enhanced leg detector based on density-weighted support vector data description (DW-SVDD) to locate the target person in a crowded environment with obstacles and occlusions. The proposed DW-SVDD leg detector offers the following benefits over the leg detectors discussed above:

– In this paper, we propose a novel approach called Density-Weighted Support Vector Data Description (DW-SVDD) that leverages 2D LiDAR data to create a 3D density representation, enabling robust leg detection in complex scenarios. The proposed DW-SVDD leg detector inputs girth G_i, width W_i characteristics and density weight $\rho(x_i)$ into support vector data description. As a result, it is possible to obtain dynamic leg data with a more reliable detecting performance, even in situations with similar obstacles and partial occlusions.
– The density weight parameter for the density distribution of LiDAR data is calculated by comparing the k-NN distance of each data point with the maximum k-NN distance of the dataset. One of the main advantages of the k-NN algorithm is its simplicity and intuitive nature. It does not assume any underlying statistical distribution of the data and can be applied to both numerical and categorical data. Therefore, our suggested DW-SVDD leg detector may be incorporated into the robot platform's low-cost embedded system.

The rest of this paper is organized as follows. Section 2 is about the DW-SVDD leg detector proposed in this paper and its application for socially assistive robots. Different experiments are designed to evaluate the traditional SVDD and DW-SVDD in Sect. 3. In this section, our proposed DW-SVDD leg detector is compared with other leg detectors in a real-time robot platform. Also, indoor and outdoor datasets are self-collected for real-time evaluation of the DW-SVDD leg detector. Additionally, all leg detectors are evaluated in CLEAR MOT metrics for multi-object tracking. Finally, Sect. 4 concludes the paper.

2 DW-SVDD Leg Detector

Support Vector Data Description (SVDD) is a powerful machine learning technique widely used for anomaly detection, novelty detection, and one-class classification tasks [18]. As an improved version of the support vector machine, SVDD generates a minimum-volume hypersphere that encircles the positive samples in feature space and identifies abnormalities of leg characteristics. By giving the density weight of each data point during the search process, the description favors data points in high-density areas. Eventually the optimal description shifts toward these dense regions. Therefore, the high-level confidence of leg data is described as a positive cluster with similar obstacles and partial occlusions.

2.1 Density Weighted Support Vector Data Description

Given the 2D range data as $x_i = \{x_1, x_2, \ldots, x_m\}$, where $x_i \in \mathbf{R}^{\mathrm{m}}$, $i = 1, 2, \ldots, m$. In the proposed DW-SVDD human leg detector, the density weight $\rho(x_i)$ depends on the k-nearest neighbor (k-NN) distance [19–21], which denotes $d\left(x_i, x_i^k\right)$ as the distance between x_i and the k-th nearest neighbor of x_i^k.

In k-NN, a query x_i^k is labeled by a majority vote y_i^k of its k-nearest neighbors in the training set. To weigh close neighbors more heavily, Fig. 2 is arranged in

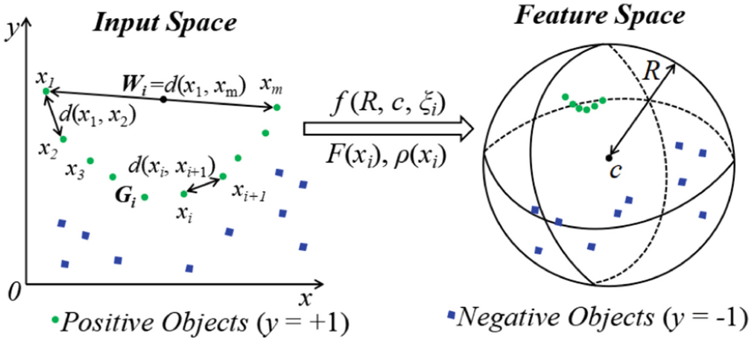


Fig. 2. Laser point in 2D range data is detected by using the DW-SVDD leg detector. DW-SVDD transforms 2D laser data into 3D density relationships, making it more suitable for leg detection in multi-person scenarios. Leg data description depends on $F(x_i)$ and $\rho(x_i)$.

increasing order in terms of Euclidean distance $d\left(x_i, x_i^k\right)$, such as:

$$\begin{cases} d\left(x_i, x_i^k\right) = \sqrt{\left(x_i - x_i^k\right)^T \left(x_i - x_i^k\right)} \\ y_i = argmax \sum_{\left(x_i^k, y_i^k \in T\right)} \delta\left(y = y_i^k\right) \end{cases} \quad (1)$$

where x_i is the training vector in the feature space, y_i is the corresponding class label in the training set $T = \{x_i, y_i\}_{i=1}^N$, x_i^k is the query for the i-th nearest neighbor among its k nearest neighbors, y_i^k is the class label for the i-th nearest neighbor among its k nearest neighbors. Also, $\delta\left(y = y_i^k\right)$ is the Dirac delta function [13].

As shown in Fig. 2, by measuring the k-NN distance, density weight $\rho(x_i)$ is set as:

$$\rho\left(x_i\right) = 1 - \frac{d\left(x_i, x_i^k\right)}{d\left(x_j, x_j^k\right)} \quad (2)$$

where $d\left(x_j, x_j^k\right)$ is the maximum k-NN distance of the dataset.

Density weight $\rho\left(x_i\right)$ measures the relative density based on the density distribution of the target data by comparing the k-NN distance of each data point with the maximum k-NN distance of the dataset. Moreover, density weight falls within the range $\rho\left(x_i\right) \in (0, 1)$.

To define attributes of legs, girth G_i, width characteristics W_i and density weight $\rho\left(x_i\right)$ are utilized to build the proposed DW-SVDD-based leg detector. The DW-SVDD optimization objective is to seek a tight data description:

$$\begin{cases} \min L\left(R, c, \xi_i\right) = R^2 + C \sum_{i=1}^n \rho\left(x_i\right) F\left(x_i\right) \xi_i \\ s.t. \left\|x_i - c\right\| \le R^2 + \xi_i, \xi_i \ge 0, F\left(x_i\right) = \left(G_i, W_i\right) \\ G_i = \sum_{j=1}^{m-1} d\left(x_j, x_{j+1}\right), W_i = d\left(x_1, x_m\right) \end{cases} \quad (3)$$

where R is the hypersphere radius, c is the center of the sphere, ξ_i is the relaxation variable for reducing the influence of singular points, and C is the penalty factor that weighs the hypersphere volume and misclassification rate. $F(x_i)$ is a function that maps data from the original space to the feature space through nonlinear transformation.

To describe the LiDAR data, the Lagrange function can be further derived as:

$$\begin{aligned} L\left(R, c, \xi_i\right) = R^2 + C \sum\nolimits_{i=1}^{N} \rho\left(x_i\right) \xi_i \\ - \sum\nolimits_{i=1}^{N} \alpha_i \left(R^2 + \xi_i - \left\| F\left(x_i\right) - c \right\|^2\right) - \sum\nolimits_{i=1}^{N} \eta_i \xi_i \end{aligned} \tag{4}$$

where α_i and η_i are the Lagrange multipliers.

The original training data are not spherically distributed for the LiDAR data that a nonlinear data relationship exists. Accordingly, anomalies cannot be isolated effectively by a hypersphere. Since partial differentiation should equal zero [13], the new equation and constraints are denoted as:

$$\begin{cases} L = \sum\nolimits_{i=1}^{m} \alpha_i k\left(x_i, x_i\right) \\ - \sum\nolimits_{i=1}^{m} \sum\nolimits_{j=1}^{m} \alpha_i \alpha_j k\left(x_i, x_j\right) \\ s.t. 0 \leq \alpha_i \leq C, \sum\nolimits_{i=1}^{m} \alpha_i = 1 \end{cases} \tag{5}$$

where $k\left(x_i, x_i\right)$ is an inner product, and $k\left(x_i, x_j\right)$ is the Kernel function.

By seeking the solution of Eq. (3), we could get the center c and radius R of the hypersphere:

$$\begin{cases} c = \sum\nolimits_{i} \alpha_i x_i \\ R = \left| x_k - c \right| \end{cases} \tag{6}$$

where x_k is the support vector, whose distance to the center of the hypersphere is the radius R.

2.2 DW-SVDD Leg Detection Algorithm

To find and save the candidate leg cluster C_k, Algorithm 1 is proposed to describe the leg detection scheme. The DW-SVDD leg detection algorithm not only takes geometric features (*e.g.*, the girth G_i and width W_i) into consideration, but also the density weight $\rho(x_i)$. As the algorithm should be executed in real-time on a socially assistive robot, short computing time and high recognition accuracy are the advantages of the proposed Algorithm 1. According to the training 2D range data x_i input, candidate leg clusters are obtained with the proposed leg detection algorithm. Besides, the central position of the LiDAR is the position of the hypersphere center c in Eq. (6).

In this paper, we present a novel leg detector for pedestrian tracking based on DW-SVDD (Density-Weighted Support Vector Data Description). The proposed approach leverages the benefits of DW-SVDD to accurately and robustly detect and distinguish human legs in complex and crowded environments. By transforming 2D laser data into 3D density relationships, our leg detector exhibits enhanced performance in multi-person scenarios, where multiple legs are present within the sensing range.

Algorithm 1: DW-SVDD Leg Detection Algorithm

Input: x_i: the training 2D range data
Output: C_k: the candidate leg cluster
for j = *1 to N* **do**
    $d\left(x_i, x_i^k\right) \rightarrow$ Eq.(1);
    if *a query* $d\left(x_i, x_i^k\right)$ *is labeled by a majority vote* $d\left(y_i, x_i^k\right)$ **then**
        $x_i^k = x_{sorted(i)}, y_i^k = y_{sorted(i)}$;
    end
end
foreach k = *1 to N* **do**
    if $x_i > x_{i+1}$ **then**
        [index, dist] = sort $\left[d\left(x_i, x_i^k\right)\right.$, ascend];
    end
end
for i = *1 to k* **do**
    $y_i \rightarrow$ Eq.(1);
    $\rho(x_i) \rightarrow$ Eq.(2);
    $G_i, W_i, F(x) \rightarrow$ Eq.(3);
    $k\left(x_i, x_j\right), L \rightarrow$ Eq.(5);
    $c, R \rightarrow$ Eq.(6);
    if $d\left(c, x_i^k\right) \in (0, \mathrm{R})$ **then**
        $C_K \rightarrow$ Save;
    end
end

2.3 Socially Assistive in Multiplayer Scenarios

As shown in Fig. 3, a socially assistive robot equipped with LiDAR, which collects the clustering of multiple pedestrians. To distinguish the tracking target $C_k\left(x_k, y_k, \theta_k\right)$ from interfering target $C_{k+1}\left(x_{k+1}, y_{k+1}, \theta_{k+1}\right)$, the robot locks the tracking target by updating the relative change angle $\Delta\theta < \delta_c$ between the robot and the followed target at each time interval, and the moving distance $L_p < \delta_f$ of the following target in real-time.

Algorithm 2: Socially Assistive Algorithm

Input: C_k: the candidate leg cluster

Output: v_l, v_r: speed input to the left and right wheel motor

for k = *1 to N* **do**

  min $d(c, C_k)$; **if** $\Delta\theta < \delta_c, L_p < \delta_f$ **then**

    $a_v, a_\omega, v_l, v_r \rightarrow$ Eq.(9);

  end

end

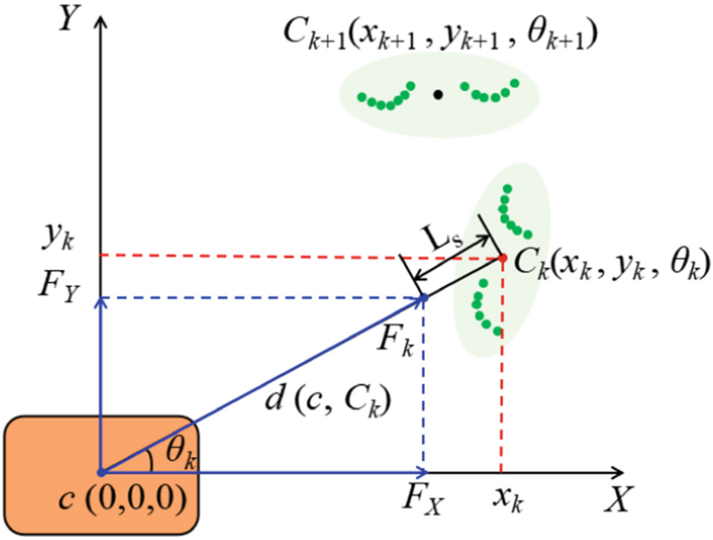


Fig. 3. Motion control in the robot coordinate system. The LiDAR is installed on the center of the robot for simplifying the coordinate system, which is set as c (0, 0, 0).

The relative angle variation $\Delta\theta$ between the following target and the robot in time interval Δt is always within the range of $\Delta\theta < \delta_c$, which is described as:

$$\begin{cases} \theta = \arctan\left(\frac{y_{k+1}-y_k}{y_{k+1}-y_k}\right) \\ \Delta\theta = \frac{(\theta_{k+\Delta t}-\theta_k)}{\Delta t} \end{cases} \tag{7}$$

The distance with the following target in Δt is always within the range of $L_p < \delta_f$, which is described as:

$$L_p = \sqrt{(x_{k+\Delta t} - x_k)^2 + (y_{k+\Delta t} - y_k)^2} \tag{8}$$

According to the dynamic model of socially assistive robots [17], the dynamic equation of the robot in the X-axis and Y-axis is denoted as:

$$\begin{cases} F_X = F_k \cos\theta_k = m a_v \\ F_Y = F_k \sin\theta_k = \frac{J a_\omega}{(d(c, C_k) - L_s)} \end{cases} \tag{9}$$

where F_X and F_Y are the component of the force on the X-axis and Y-axis, respectively. F_k is the force that pulls the robot to follow the target. m is the weight of the robot, J is the rotational inertia of the robot, a_v and a_ω are linear acceleration and angular acceleration, respectively. L_s is the safe distance that the robot would not collide with the following target.

Algorithm 2 defines how the socially assistive robot maintains a safe distance L_s between the robot and the following target. More than one cluster may be available in a cluttered environment [22]. For the safety of the following target, the robot should follow the nearest cluster.

3 Self-collected Dataset

The majority of datasets are only able to gather 2D range information while a single target is moving. Despite the fact that Beyer *et al.* [12] made their LiDAR dataset-a 10-h record for an aged care facility-available for download. The users probably live in a relatively clean area because there aren't many impediments in this dataset. Because labelling is so expensive, it is hard to discover datasets with several people or obstacles. Additionally, data from the Leon@Home Testbed [25], which includes location estimates determined by two human trackers, has been gathered to assess service robots in a realistic home context. Obviously, the aforementioned data is insufficient to train the suggested detector.

3.1 Overview

Table 1 contains self-collected OpenField and MessyIndoor datasets for comparing and evaluating the DW-SVDD leg detector. Additionally, because our team is dedicated to using open-source datasets, this data will be posted on the GitHub website along with any supporting materials, such as movies and rosbag data.

Table 1. Overview of Self-Collected Datasets

Dataset Topics	OpenField	MessyIndoor
Running Time	46.2 s	38.2 s
Humans Number	5	5
Obstacles Number	0	5
Topics: /scan	933 msgs	898 msgs
Topics: /odom	469 msgs	449 msgs
Topics: /leg_clusters	926 msgs	885 msgs

3.2 Training Dataset

Experiments in a multi-person setting are split into detection experiments and following experiments in order to validate the DW-SVDD approach for socially assistive tasks. As shown in Fig. 4, the man wearing a white shirt is the target being followed, and the man wearing a black shirt is the target being interfered with. As shown in Figs. 4a and 4b, the interfering target frequently creates a barrier between the robot and the target being pursued in a multi-person environment. When legs are too close together, the mobile robot interprets the grouping of several legs as an incorrect goal and sets misleading data (Fig. 4c). A control group of real data was also established (Fig. 4d).

4 Experiments of Leg Detection

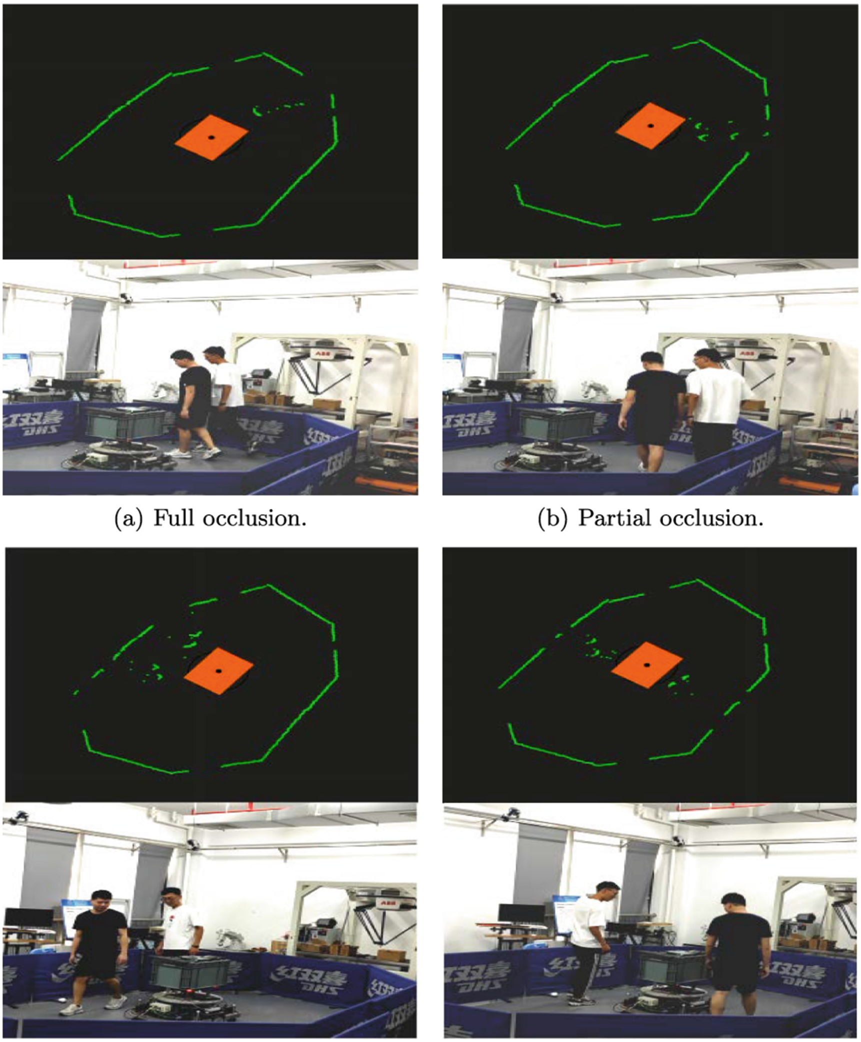
4.1 SVDD Leg Detector vs DW-SVDD Leg Detector

SVDD and DW-SVDD leg detectors are evaluated on the Redmi laptop computer, which is equipped with an Intel Core i5-10300H processor, 16 GB RAM, an Nvidia GTX 1650 GPU, and an Ubuntu 20.04 operating system. All the contrastive methods are implemented in the ROS (Robot Operating System) and do not employ any parallel acceleration optimization techniques.

The data in 3200 laser data samples was discovered to objectively examine the recognition and detection of targets by mobile robots in a multi-person scene. The training set from Fig. 4 is used to test the SVDD leg detector in Fig. 5. It is obvious that the SVDD method's excessive reliance on leg characteristics is the reason why the detection results are inadequate.

Figure 5(a) shows that the robot only detects one human leg information because the followed target is completely blocked. The robot only detects part of the human leg information in Fig. 5(b). In Fig. 5(c), the distance between the followed target and the interfering target is very close, detectors have difficulty clustering these features into the following target. The control group in Fig. 5(d) also demonstrated successful detection.

Figure 6 demonstrates how much more accurate and efficient our suggested strategy is. The identification of laser point sequences in close proximity is made easier since density information is taken into account. Figure 6(a) demonstrates that the robot only picks up information about one human leg because all other target information was lost. The DW-SVDD approach in Fig. 6(c) obviously outperforms the SVDD method since it recognizes two targets as opposed to one in Fig. 5(c). Additionally, Fig. 6(d) more clearly illustrates the test's outcome.



(a) Full occlusion. (b) Partial occlusion.

(c) Walking one behind the other. (d) Walking and splitting up.

Fig. 4. Training dataset: In a setting with multiple people, following and interfering targets are moving about at various positions.

The quantitative evaluation results are shown in Table 2. We could conclude that DW-SVDD leg detector has advantages over SVDD-based leg detector in terms of multi-target classification. As for Area Under the Curve (AUC), the

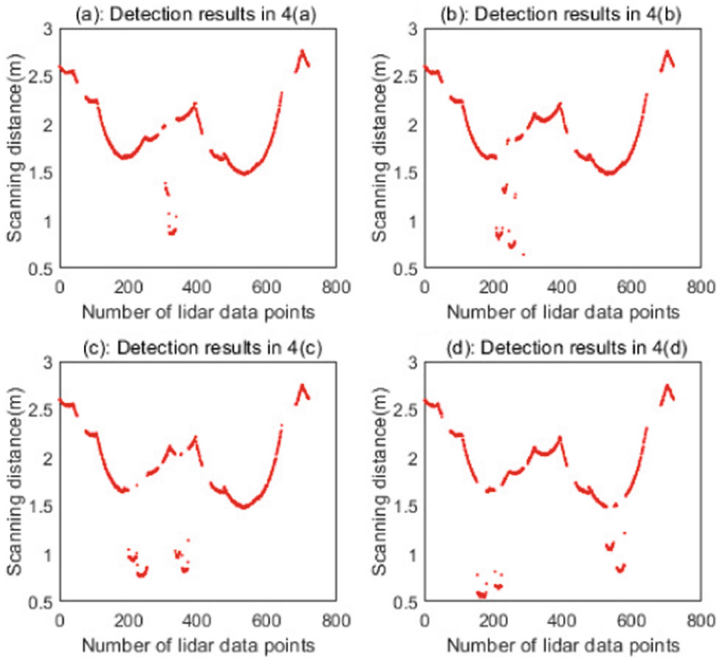


Fig. 5. The detection results of the SVDD leg detector for the four scenarios in Fig. 4.

Table 2. Comparing the Detection Results of SVDD and DW-SVDD Detectors

Leg Detector	SVDD [14,17]	DW-SVDD
Running Time	41.2 ms	47.0 ms
Number of Iterations	9	15
Average Lost Points	15	6
AUC	60.36%	73.41%
Variance of Distance and Radius (s_{d-r}^2)	19.21	12.30

DW-SVDD leg detector is 73.41%, which shows the higher accuracy of the detector. To describe the variance between the distance and radius, s_{d-r}^2 is defined to reflect the fluctuation of data:

$$s_{d-r}^2 = \frac{\sum_{i=1}^{n} \left(d\left(x_i, x_i^k\right) - r\right)^2}{n} \tag{10}$$

where the distance $d\left(x_i, x_i^k\right)$ and the radius r are described in Algorithm 1 and Eq. (6), respectively.

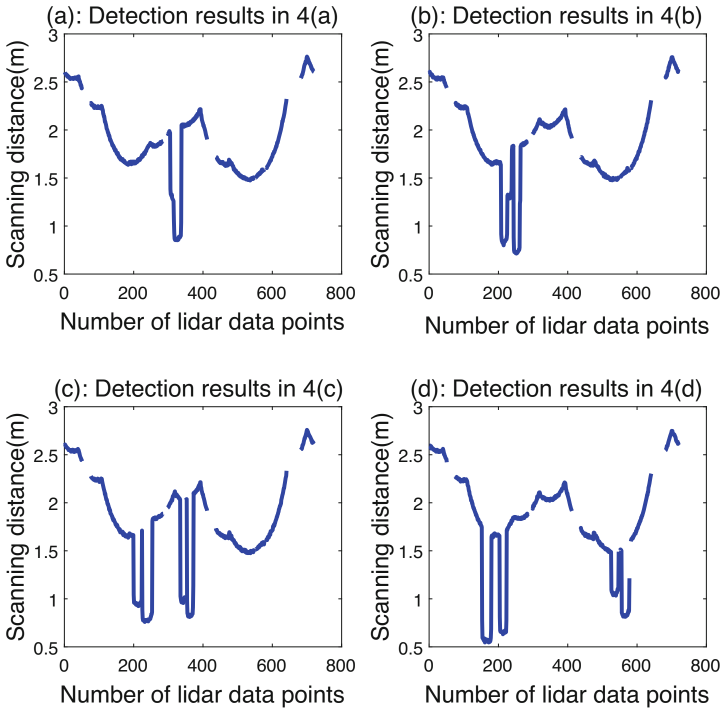


Fig. 6. The detection results of the DW-SVDD leg detector for the four scenarios in Fig. 4.

4.2 DW-SVDD Leg Detector in Real-Time Robot Platform

This detector is practically implemented with a 2-wheeled mobile robot in order to test the effectiveness of the DW-SVDD-based leg detector in a real-world setting. To gather 2D range data, a LiDAR with a 360° measurement angle (R2000, Pepperl+Fuchs GmbH, Germany) was put in the middle of the mobile robot. It has a 360° scanning angle and a measurement range of up to 30 m. With a measurement speed of 250 kHz and a scanning rate of 10 m/s, laser data on the leg can be detected.

Leg Detection for Partial Occlusions. One interference target was intended to pass through the target person and the robot in Fig. 7. The detector can easily be shown to discriminate between the two targets despite the interference target and the target after it being so close together that they are even partially obscured. It should be noted that if a person's legs are totally hidden, they cannot be recognized. In this situation, the LiDAR is deprived of all knowledge on the features of the human leg.

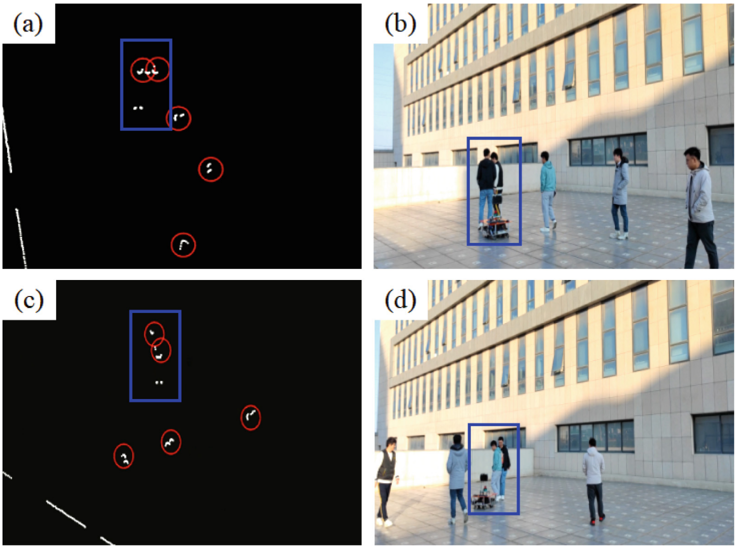


Fig. 7. The DW-SVDD leg detector could be used by socially assistive robots to identify partial occlusions. (a)(b) In blue box, the DW-SVDD leg detector could be able to recognize both people in real-time if they are standing adjacent to the target. (c)(d) The DW-SVDD leg detector also functions when someone passes by the target in blue box. (Color figure online)

Leg Detection with Similar Obstacles. In Fig. 8, despite the close proximity of the interference target and the target, even partially obscuring the target, the detector can be seen clearly differentiating between the two targets. The density weighting, which considers the distance between the legs and attempts to separate legs with similar obstacles, makes this detection successful.

Table 3. Evaluation Results of Three Leg Detectors

Leg Detector	LegDetector [24]	SVDD [14,17]	DW-SVDD
ID Switch	115	137	155
Miss	2371 msgs	2113 msgs	1474 msgs
FP	2294 msgs	2110 msgs	1890 msgs
MOTA	18.70%	27.41%	34.96%
MOTP	26.97%	37.02%	39.17%

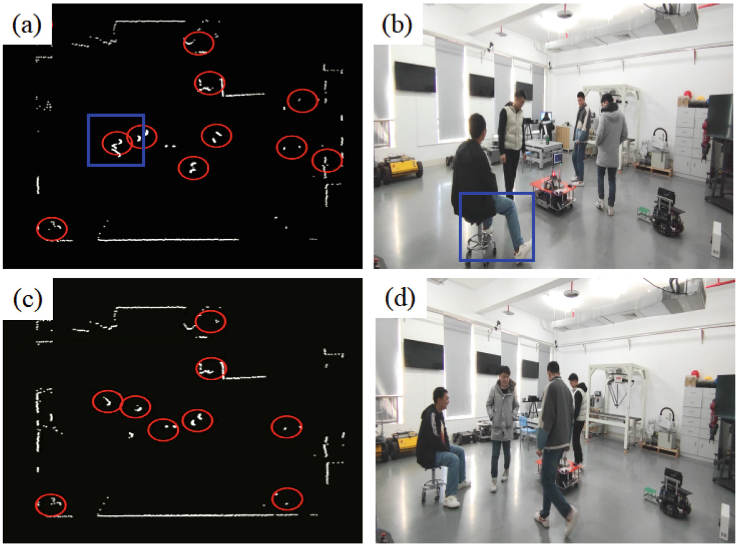


Fig. 8. The DW-SVDD leg detector could be used by socially assistive robots to identify similar obstacles. (a)(b) In the blue box, the DW-SVDD leg detector might be able to recognize both people in real-time if they are seated in similar obstacles. (c)(d) MessyIndoor Dataset detection results. (Color figure online)

4.3 Leg Detection in CLEAR MOT Metrics

Due to a scarcity of publicly available code, we could only find an open-source ROS-enabled leg detector package [23], which is based on the geometric constraints of legs.

The CLEAR MOT metrics [24] are adopted for the quantitative evaluation of socially assistive robots. These metrics are commonly utilized for multi-object tracking, and provide scorings of ID switches (ID_k), misses $(Miss_k)$, false positives (FP_k), valid assignments, and precision of matchings cumulated from every frame. These scores can be aggregated in multi-object tracking accuracy (MOTA) score and multi-object tracking precision (MOTP):

$$\begin{cases} MOTA = 1 - \frac{\sum_k (I_k + Miss_k + FP_k)}{\sum_k g_k} \\ MOTP = \frac{\sum_{i,k} d(x_i, x_k)}{\sum_k c_k} \end{cases} \tag{11}$$

where g_k is the ground truth annotations, c_k is the number of matchings between the estimated person and ground truth positions at time k. The distance $d(x_i, x_k)$ is described in Algorithm 1.

As shown in Table 3, results indicate that the MOTA and MOTP of the proposed DW-SVDD leg detector achieved 34.96% and 39.17%, respectively.

5 Conclusion

This study presents a DW-SVDD leg detector to detect, track, and follow human targets using a LiDAR at leg height in a socially assistive robot. To improve the leg detection performance for partial occlusions and similar obstacles in 2D range data, the proposed leg detector integrates k-NN distance, density weight, and support vector data description. Experimental results show that the DW-SVDD leg detectoroutperforms the other two detectors, exhibiting superior clustering and classification performance in a cluttered environment. Specifically, the proposed leg detector achieved 34.96% and 39.17% for MOTA and MOTP, respectively. However, there are limitations to the detector, as it falsely detects similar obstacles in Fig. 8 due to not filtering out noise from distant laser points. In addition, this leg detector fails to detect the target person who wears a long skirt.

In future works, we plan to develop a stable and dependable socially assistive system for mobile robots. This will involve using a high-performance robot controller to test out a sensor fusion strategy, which will include running a lightweight neural network-based detector. Overall, the DW-SVDD leg detector shows promising results and has the potential to be integrated into a more comprehensive socially assistive scenario in the future.

References

1. Hu, Y., Abe, N., Benallegue, M., et al.: Toward active physical human-robot interaction: quantifying the human state during interactions. IEEE Trans. Hum. Mach. Syst. **52**(3), 367–378 (2022)
2. Yuan, J., Zhang, S., Sun, Q., et al.: Laser-based intersection-aware human following with a mobile robot in indoor environments. IEEE Trans. Syst. Man Cybern. Syst. **51**(1), 354–369 (2018)
3. Xue, G., Yao, H., Zhang, Y., et al.: UWB-based adaptable side-by-side following for human-following robots. In: 2022 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 333–338, IEEE, Jinghong, China (2022)
4. Liu, P., Yao, H., Dai, H., et al.: The detection and following of human legs based on feature optimized HDBSCAN for mobile robot. J. Phys: Conf. Ser. **2216**(1), 012009 (2022)
5. Zhou, G., Zhao, R., Yao, H., et al.: A low-cost conductive-textile based multifunctional flexible capacitive sensor for human motion tracking. In: 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 317–321, IEEE, Sanya, China (2021)
6. Su, S., Cheng, S., Dai, H., et al.: An efficient human-following method by fusing kernelized correlation filter and depth information for mobile robot. In: 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 2099–2104, IEEE, Dali, China (2019)
7. EffiBOT (AGV/AMR). https://www.effidence.com/en/effibot/. Accessed 1 Oct 2022
8. Li, H., Dong, Y., Li, X.: Object-aware bounding box regression for online multi-object tracking. Neurocomputing **518**, 440–452 (2023)

9. Nurunnabi, A., Sadahiro, Y., Laefer, D.F.: Robust statistical approaches for circle fitting in laser scanning three-dimensional point cloud data. Pattern Recogn. **81**, 417–431 (2018)
10. Arras, K.O., Mozos, O.M., Burgard, W.: Using boosted features for the detection of people in 2D range data. In: Proceedings 2007 IEEE International Conference on Robotics and Automation, pp. 3402–3407, IEEE, Rome, Italy (2007)
11. Li, D., Li, L., Li, Y., et al.: A multi-type features method for leg detection in 2-D laser range data. IEEE Sens. J. **18**(4), 1675–1684 (2017)
12. Beyer, L., Hermans, A., Linder, T., et al.: Deep person detection in two-dimensional range data. IEEE Robot. Autom. Lett. **3**(3), 2726–2733 (2018)
13. Tax, D.M., Duin, R.P.: Support vector domain description. Pattern Recogn. Lett. **20**(11–13), 1191–1199 (1999)
14. Chung, W., Kim, H., Yoo, Y., et al.: The detection and following of human legs through inductive approaches for a mobile robot with a single laser range finder. IEEE Trans. Industr. Electron. **59**(8), 3156–3166 (2011)
15. Cha, D., Chung, W.: Human-leg detection in 3D feature space for a person-following mobile robot using 2D LiDARs. Int. J. Precis. Eng. Manuf. **21**, 1299–1307 (2020)
16. Jung, E.J., Lee, J.H., Yi, B.J., et al.: Development of a laser-range-finder-based human tracking and control algorithm for a marathoner service robot. IEEE/ASME Trans. Mechatron. **19**(6), 1963–1976 (2013)
17. Yao, H., Dai, H., Zhao, E., et al.: Laser-based side-by-side following for human-following robots. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2651–2656, IEEE, Prague, Czech Republic (2021)
18. Nguyen, X.T., Kim, H., Lee, H.J.: An efficient sampling algorithm with a K-NN expanding operator for depth data acquisition in a LiDAR system. IEEE Trans. Circuits Syst. Video Technol. **30**(12), 4700–4714 (2020)
19. Cha, M., Kim, J.S., Baek, J.G.: Density weighted support vector data description. Expert Syst. Appl. **41**(7), 3343–3350 (2014)
20. Wu, X., Liu, S., Bai, Y.: The manifold regularized SVDD for noisy label detection. Inf. Sci. **619**, 235–248 (2023)
21. Sung, Y., Chung, W.: Hierarchical sample-based joint probabilistic data association filter for following human legs using a mobile robot in a cluttered environment. IEEE Trans. Hum. Mach. Syst. **46**(3), 340–349 (2015)
22. Standard Classification (Banana Dataset). https://www.kaggle.com/saranchandar/standard-classification-banana-dataset. Accessed 15 Sep 2018
23. ROS Leg_Detector. http://wiki.ros.org/leg_detector. Accessed 10 Sep 2017
24. Bernardin, K., Stiefelhagen, R.: Evaluating multiple object tracking performance: the clear mot metrics. EURASIP J. Image Video Process. **2008**, 1–10 (2008)
25. Álvarez-Aparicio, C., Guerrero-Higueras, Á.M., Olivera, M.C.C.: Benchmark dataset for evaluation of range-based people tracker classifiers in mobile robots. Front. Neurorobot. **11**, 72 (2017). https://www.frontiersin.org/articles/10.3389/fnbot.2017.00072/full