A REAL-TIME IMPROVED PEDESTRIAN DEAD RECKONING TRAJECTORY TRACKING ALGORITHM

Wang Xuerui 1, Lv Jingguo 2, Yun Zehui 3

¹ School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China,1742208950@qq.com;

² Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing 100044, China, lvjg@ bucea.edu.cn;

³ School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China,463833740@qq.com.

Commission III, WG III/1

KEY WORDS: Pedestrian positioning, Target detection, Target tracking, Distance detection

ABSTRACT:

Accurate location of pedestrians plays a crucial role in emergency relief, traffic control, crowd behavior analysis and other aspects. Especially in the context of the COVID-19 pandemic in recent years, pedestrian location technology can help relevant departments to complete target screening more quickly. However, the Pedestrian Dead Reckoning algorithm can only calculate the target trajectory through the sensor return value, but can not carry out real-time trajectory correction and location. With the rapid development of deep learning, object detection and tracking technology based on computer vision has been applied to pedestrian location, but there are two challenges in the application process. Firstly, in the pedestrian gathering scene, the target number base is large, so the accuracy of the current detection algorithm needs to be improved and the model drift index of the tracking algorithm needs to be reduced. Secondly, there is a certain distortion between the real three-dimensional coordinate space of pedestrians and the two-dimensional image captured by the camera, and the transformation of the spatial coordinate of the target point is a technical difficulty.In this regard, first of all, to improve the accuracy of pedestrian target detection in crowded scenes, this paper adopts the method of improving the generalization of network to pedestrian target, and uses k-means algorithm to find the best prior frame of pedestrian, and sets the width to height ratio suitable for the target. Secondly, to solve the problem of model drift in the above tracking process, this paper proposes a binary classification model based on target appearance difference, which introduces target context information as a new target distinguishing feature when two or more targets are similar. Finally, in order to obtain more accurate coordinate position information, this paper combines the inverse perspective algorithm to calculate the target coordinates into the coordinates in the world coordinate system, and calculates the exact position of the target in the aerial view, as well as the distance between the targets or the current flow of people. In order to evaluate the effectiveness of the proposed algorithm, experiments on target detection, tracking and precise positioning were carried out in different intensity scenarios to verify the feasibility of the proposed method.

1. INTRODUCTION

With the growth of people's demand for location information, global navigation and positioning technology has become increasingly mature in outdoor navigation and positioning and has been successfully applied to various service software. However, indoor positioning is still developing due to the complexity of the environment and the high precision required to predict the target trajectory.PDR(Pedestrian Dead Reckoning) is also a kind of indoor positioning technology, which uses the acceleration, gyroscope and other sensor data collected during the Pedestrian movement to judge the current state of the Pedestrian, and calculates the step length and course of each step during the Pedestrian movement, so as to determine the current position of the Pedestrian coordinates. However, the biggest disadvantage of pedestrian navigation prediction is the accumulation of trajectory errors. The N+1 coordinate calculated is easily affected by the errors of the first N coordinates, resulting in large deviations of the later coordinates. Therefore, in order to solve the above problems, other information is usually added to the PDR algorithm to improve the positioning accuracy of the algorithm and achieve real-time target positioning. There are three improved pedestrian estimation and location methods(Zhang G et al.,2019; Klein I et al.,2018; Yuqin et al.,2019). The first is to

know the moving area of the target, display the calculated point coordinates on the map through the projection matching method on the basis of PDR, and improve the point accuracy through the passable path of the map. The second is the PDR positioning algorithm based on particle filter. The accelerometer built in the mobile phone is used for step detection and step measurement, the gyroscope and magnetometer are used for direction estimation, and the particle filter algorithm is used to filter and fuse the positioning data in combination with the map matching information, so as to improve the positioning accuracy of pedestrian dead reckoning. The third is the indoor three-dimensional positioning algorithm based on multi-sensor fusion. Because different sensors have different disadvantages and advantages, how to use each sensor efficiently is also a key research to improve the level of the algorithm. On the basis of PDR, this paper adds computer vision target detection and target tracking technology, aiming to improve the dead dead position of pedestrian in real time and reduce the trajectory error.At the same time, thanks to the addition of computer vision technology, it can also calculate the flow of people in the current location and the distance between people.

As an important research content in computer vision, target detection and target tracking have developed many algorithms up to now.RCNN, as a classical algorithm for deep learning target detection, has developed fast-RCNN, faster-RCNN and other algorithms based on it(GIRSHICK R et al., 2014; ROSS G,2015;REN S et al.,2017). These algorithms have high detection accuracy and have attracted the attention of a large number of scholars.Since then, YOLO series(Redmon J et al.,2016;Redmon T et al.,2017;Redmon J al.,2018;BOCHKOVSKIY A et al.,2020), SSD (Liu W et al.,2016) and other algorithms have appeared successively. Such algorithms have improved detection speed, made realtime detection possible, and improved performance in all aspects.SORT(Bewley A et al.,2016) algorithm is an online real-time multi-target tracking algorithm, which uses Kalman filter to predict the state and implements a fast multi-target tracking result based on the association between the detection frame and the Hungarian assignment of the IOU of the prediction frame.Wojke N proposed DeepSORT(Wojke N et al.,2017) algorithm on the basis of SORT, adding cascade algorithm to reduce ID Switch and pedestrian re-recognition model. Training by combining motion and appearance information in order to reduce the influence of occlusion. Although these methods have achieved good results, the existing methods still need to be improved due to the complex detection background, different human poses and different degrees of occlusion after the combination of pedestrian prediction technology.

In order to reduce the pedestrian navigation position error accumulation in the process of calculation problem of target tracking accuracy and the effect of this article in the PDR algorithm fusion target detection and target tracking technology in computer vision and join the inverse perspective mapping algorithm, to two-dimensional coordinate transformation of a video to the world coordinate system of coordinates, making them in motion in the process of automatic correction trajectory error.In order to solve the problem of detection accuracy and tracking model drift in target detection and target tracking, this paper improves the generalization of detection network to pedestrian targets, and proposes a binary classification model based on the difference of target appearance that is more likely to distinguish the features of similar targets.

2. METHODOLOGY

2.1 Pedestrian Dead Reckoning

The basic principle of PDR algorithm is to estimate the next position coordinate of the target by using the azimuth Angle $\theta 1$

and step size d1 obtained by the sensor when the initial position of the target is known (X_0, y_0) .Such repeated calculation can obtain the position information reached by the Nth step of the target (X_n, y_n) .The coordinate calculation is shown in Formula 1, and the basic model of the algorithm is shown in Figure 1.

$$\begin{cases} x_k = x_0 + \sum_{i=1}^k d_i \sin \theta_i \\ y_k = y_0 + \sum_{i=1}^k d_i \cos \theta_i \end{cases}$$
(1)

The azimuth Angle $\theta 1$ and step size d_1 in the figure mainly consist of three elements: gait detection, step size estimation and direction estimation.Gait detection methods include time-domain analysis and frequency-domain analysis. Generally, empirical model is used to calculate step size. Direction estimation is obtained by digital compass, gyroscope, various sensors, etc.



Figure 1. Basic model of PDR algorithm.

2.2 Target Detection

The tracking algorithm in this paper belongs to the tracking-bydetection method, and the matching between the detection frame and the prediction frame is carried out on the basis of detection. Therefore, the target detection algorithm directly affects the statistics result of people flow.As the research purpose of this paper is related to real-time detection, yOLOV4 algorithm is selected in this paper, and the algorithm network framework is shown in Fig2.The algorithm performs well in the field of target detection with high accuracy and high speed.



Figure 2. YOLOv4.

As the research purpose of this paper is to count the flow of people, the detection scenes are generally crowded with targets, and the non-maximum inhibition threshold of other detection algorithms is set strictly, which results in the suppression of targets that are very close to each other. However, yOLOV4 algorithm changes the detection NMS algorithm(Neubeck, A et al.,2006) to soft-NMS algorithm(Bodla N et al.,2017). This algorithm reduces the suppression effect of targets with similar distance without increasing the amount of computation and is more suitable for the calculation of human flow.At the same time, the following improvements are made to the target missed detection network:

(1)K-means method was used to find the best prior frame on the training set, and the width-to-height ratio suitable for human body was set (1:2) * (64,128,256) to reduce redundancy. (2)The self-made data set was used to train the network, and the training data in crowded scenarios were added to improve the generalization of the model.

2.2.1 Soft - NMS algorithm : The NMS algorithm is used to remove duplicate boxes in YOLOV4. Relative to a target, the generated box with the highest confidence is selected to be reserved as the target box of the target. At the same time, the operation of NMS exists the situation that only one target box is reserved for two similar targets. However, in the crowded crowd scenarios studied in this paper, most of the targets are very similar, and the NMS algorithm is obviously not suitable for the detection algorithm used in this paper. Soft-nms algorithm is an improvement of NMS algorithm. It does not delete all the boxes whose IOU is greater than the threshold value, but reduces their confidence and sorts the confidence of the target box, and then selects the threshold value to get the final target box. The algorithm flow chart is shown in Figure 3:



Figure 3. Flow chart of soft-NMS algorithm.

Input the initial detection frame stored in part B set; The detection score value of each detection frame response is stored in S set; Is the set NMS threshold.

2.2.2 YOLOV4 Target width and height setting : Yolov4 algorithm adopts the method of YOLOv3 in multi-scale prediction, fuses the feature maps of three different levels, and inherits the detection head of YOLOv3.While training on the COCO dataset, The nine Anchors chosen by YOLOv4 are (10x 13),(16x 30), (33 X 23),(30x 61),(62 X 45), (59 X 119),(116 X 90), (156 X 198), (373 X 326).Since the detection task target in this paper is only the crowd, this setting is not applicable to the target detection in this paper.Firstly, k-means method was used to cluster the target frames in the training set to find the best prior frames, and then the width-to-height ratio suitable for human body was set according to the proportion of human body (1:2) * (64,128,256). Figure 4 shows the results displayed by the improved anchor in mot16-08 data set, with sizes of 64* 128,128 * 256,256 *512 respectively. It can be seen from Figure * that the width and height values set in this paper can accurately cover the detected targets.



Figure 4. Improved Anchor size.

2.3 Target Tracking Based on Trajectory Prediction

In this paper, Deep sort tracking algorithm is used to track the contacts of the target.Deep Sort is a multi-target tracking algorithm, which detects the target for each frame of the video, and then matches the previous motion track with the current detected object through the Hungarian matching algorithm with weights to form the motion track of the object.After ID assignment is made to the results detected by YOLOv4 above, they are input into the target tracker, and the coordinates of the center point of each detected target frame will be input into the Kalman filter to predict the center point at the next moment.The Hungarian matching algorithm is used to match the target detected in the next frame with the predicted result of this frame to get the real location of the target.

2.3.1 Kalman filtering : Kalman filtering algorithm is a recursive predictive filtering algorithm, which predicts k+1 moment through the data of k moment. The algorithm consists of two steps: the first step is prediction, and the second step is correction, which are described by time update equation and state update equation respectively(Yong Chen, et al.,2020).In any given video, the target trajectory is modeled as $\mathbf{x} = [\mathbf{p}, \mathbf{v}]^T$, where P represents the position information of the target in the image coordinates, and V represents the speed information of the target in the image coordinates.

Time update equation of Kalman filter:

$$X_{(k|k-1)} = FX_{(k-1|k-1)} + AU_{(k)}$$
(2)

$$P_{(k|k-1)} = FP_{(k-1|k-1)}F^{T} + Q$$
(3)

Where, F is the change matrix coefficient of position and velocity from moment K to moment k+1, $AU_{(k)}$ is the constant term of change matrix, and Q is the uncertainty in change. Kalman filter state update equation:

$$K_{g(k)} = P_{(k|k-1)}H^{T} / (HP_{(k|k-1)}H^{T} + R)$$
(4)

$$X_{(k|k)} = X_{(k|k-1)} + K_{g(k)}(Z_{(k)} - HX_{(k|k-1)})$$
(5)

$$P_{(k|k)} = (I - K_{g(k)}H)P_{(k|k-1)}$$
(6)

Where, R is the observation error covariance matrix, H is the conversion matrix coefficient between the observation result and the state, and Z is the conversion process matrix between the observation result and the state.

2.3.2 Hungarian Algorithm : The Hungarian matching algorithm matches the predicted target box generated by the above Kalman filter with the detected target box obtained by the target detection algorithm. Through the IoU distance between the detected target and all the predicted boundary boxes of the predicted target, a matrix is formed, and the optimal matching is obtained through the matrix operation. At the same time, an IoUmin threshold is set in the algorithm,

which means that when the distance between two matching boxes is less than this threshold, the matching is rejected, reducing unnecessary work. When the target is blocked by other targets in a short period of time, the optimal solution is first found through the IOU distance, and then corrected through target detection, while the covered target will not be affected.

2.3.3 contextual information : In order to solve the ID conversion problem, this paper proposes a target box matching method based on color and HOG feature. The Deep sort tracking algorithm uses the Hungarian matching algorithm to complete the matching between the target frame of frame (T-1) and the target frame of frame T, while the Hungarian matching algorithm uses Mahalanobis distance or IOU to calculate the matching weight, When the distance between the two targets At and Bt in T frame is too close, the matching weight of these two targets will be similar or even the same as that of A(T-1) or B(t-1) target frame in (T-1) frame, leading to matching errors, reducing the discriminant ability of tracking algorithm and resulting in model drift. Therefore, by making full use of the connected information between the upper and lower frames of the target, the algorithm has the ability of context awareness, which is of great significance for tracking model to improve the model drift problem.



(b)extract HOG leatur

Figure 5. Feature extraction results.

Since the tracking target in this paper is a crowd and individual clothing is obviously different, on the basis of the above target tracking, the matching of target point color information and HOG feature is added in the target box matching to distinguish the two targets that are close to each other.Figure 5(b) and (c) are the extraction results of HOG feature and RGB color histogram in Figure 5(a). It can be seen that HOG feature statistics can represent the gradient direction histogram of the image, and target features can be well expressed in scenes such as illumination changes and plane rotation. Color features can be used to individually distinguish targets.

Different features have different characteristics. According to Formula (7) (Matthias Mueller et al.,2017), the target color histogram feature response graph and HOG feature response graph are fused with fixed weight, Where r is the fixed weight.

$$F_{\text{trad}} = r * f_{\text{hog}} + (1 - r) * f_{\text{color}}$$
(7)

On the basis of Deep Sort algorithm, we propose a simple, effective and universal context information calculation method. Adding context information will have a certain influence on trajectory matching.Generally, some tracks will not match when the target is occlude, the distance between the targets is too close or the target scale changes. Therefore, we associate the low score detection frame with these mismatched tracks to recover the objects in the low score detection frame and filter out the background at the same time. Algorithm pseudocode is shown in algorithm 1.

The algorithm input is a video sequence V, improved YOLOv4 detector, Kalman filter KF. Two thresholds and are set at the same time, where is detection threshold and is tracking threshold. The algorithm output is the target trajectory T.

2.4 Coordinate Transformation

In this paper, inverse perspective mapping algorithm is applied to coordinate transformation and distance detection of objects. The image is converted from camera coordinates to world coordinates, and the distance between objects is calculated by Euclidean distance.

Figure 6 shows the geometric relationship between the camera and the scene taken in the world coordinate system. The object in the upper left corner is the camera, O is the light center of the camera, and I is the point projected on the ground by the light center of the camera, namely the origin of the world coordinate system. In the figure, ABCD is the mapping range of the camera on the ground; Y-axis is the direction of the trapezoidal ABCD midline; X-axis is the direction perpendicular to Y-axis on the ABCD plane; z-axis is the straight line perpendicular to plane ABCD and passing through point I. Point P is any point in the camera's field of view, and the coordinates are(Xp, Yp). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVI-3/W1-2022 7th Intl. Conference on Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS 2022), 18–19 March 2022, Wuhan, China





Figure 6. The geometric relationship of the camera field in the world coordinate system.

There are mainly three simplified formulas for inverse perspective transformation derived from Yu Cao et al.(2011). This paper uses the simplified formula for inverse perspective transformation of Formula 2 (XU You-chun et al.,2003) in this literature to obtain the inverse perspective transformation of point P from the image plane coordinate system to the world coordinate system, and the expression is:

$$u = \frac{mY_p - hmtan\psi}{2Y_p tan\psi tan\alpha + 2htan\alpha}$$
(8)

$$v = \frac{nX_p}{2\tan\beta\sqrt{h^2 + Y_p^2}}\sqrt{1 + (\frac{2u\tan\alpha}{m})^2}$$
(9)

The known parameters in the formula mainly include: the height of the camera's optical center from the ground h (OI in Figure 6); The longitudinal field of view Angle of the camera is 2α , namely Angle GOE; Transverse field of view Angle 2β , namely Angle JOH; The top Angle Ψ of the camera is FOI; The resolution of the camera is MXN, where M is the vertical resolution and N is the horizontal resolution. Obtain the x coordinates of point P in the image plane. V is the y coordinate of point P in the image plane.

The camera coordinates are converted into world coordinates through the above formula transformation, and the coordinate transformation model is established. When the distance between two targets needs to be measured, the coordinates of the center point of the bottom edge of the two target frames are selected respectively. After the conversion to world coordinates, the distance between two targets can be obtained through Euclidean distance calculation method in the world coordinate system. Finally, the real distance of the two objects is calculated according to the proportion between the precalculated world coordinate system and the real distance.

2.5 Traffic Statistics Method

Traditional traffic statistics often use single-line method, that is, delimit a line in the detection area and calculate the number of people passing this line to carry out traffic statistics or directly count the number of targets on the picture. However, this method can not flexibly set the target detection area, can not accurately judge the direction of pedestrian movement and other problems. In order to solve the above problems, this paper proposes a regional statistical method. As shown in FIG. 7, an additional line is set on the basis of the traditional single-line method to circle the detection area and record the target movement direction, which makes up for the shortcomings of the single-line method, creates favorable conditions for target detection and tracking, and improves the statistical accuracy of human flow.



Figure 7. Intelligent statistics process of human flow.

2.6 Error Correction

After the visual sensor is added, the moving target is detected in the video, and the position of the target in the image is determined and converted into coordinates in space. Using target tracking technology to track the target to determine the target coordinates at time T, and then compare it with the target position calculated by PDR to complete the target error correction.

3. RESULTS AND DISCUSSION

In this paper, the CPU is Intel(R) Core(TM) i7-9700 and the memory is 16GB under the Windows10 operating system. CUDA10.1 and CUDNN7.6 are used for GPU acceleration. The experimental environment is compiled in Visual Studio 2015.

3.1 Datasets and evaluation indicators

In this experiment, the improved detection network was first trained by using a self-made data set. There were 1082 images containing pedestrian targets in the data set, which were annotated by Labeling. The data set includes multiple scenarios of pedestrian behavior, such as streets, malls, lawns, beaches, etc. In order to improve the generalization ability of strong model for crowd detection in dense scenes, we conducted experiments on Shanghai part_B_final dataset. This data set is often used for crowd counting. The collection scene is a street in Shanghai, and the crowd in each image is relatively dense.

Because the passenger flow in crowded places needs to be taken into account in the detection process of this model, the paper chooses to carry out passenger flow detection on the whole body of the target in consideration of the occlusion in this environment and the tracking of continuous frames. In order to evaluate the effectiveness of the algorithm, this paper adopts Mean Average Precision (mAP) as the evaluation standard of the target detection algorithm, and frames per second (FPS) as the real-time measurement standard of the target tracking speed. Recall, Precision and F value are used as the evaluation criteria of the traffic statistics algorithm, and the formula of each evaluation index is shown as follows.

$$R = \frac{TP}{TP + FN}$$
(10)

$$P = \frac{TP}{TP + FP}$$
(11)

$$F = \frac{2 * P * R}{P + R}$$
(12)

Where, TP refers to the number of pedestrians who are actually counted correctly, that is, the number of positive inspection; FP refers to the number of pedestrians who are actually nonpedestrians but are wrongly counted as pedestrians, that is, the number of false detection; FN refers to the actual number of pedestrians but not counted, that is, the number of missed detection.

3.2 Analysis of experimental results

3.2.1 Experimental analysis of target detection : We tested the algorithm on our own data set, and the average accuracy reached 80.10%. The improved algorithm in this paper, the original algorithm and the images of Faster R-CNN in Shanghai data set are tested. The test results are shown in Table 1, indicating that the accuracy of pedestrian detection in this algorithm is improved in dense scenes.

| Algorithm | MAP/% |
|------------------|-------|
| Faster R-CNN | 63.27 |
| Primal algorithm | 62.85 |
| Ours | 64.30 |

Table 1. Algorithm detection results.

3.2.2 Experimental analysis of target tracking : In the statistical experiment of pedestrian flow, three surveillance videos downloaded from the Internet with different intensity and different scenes were used for testing. The specific information of the test videos is shown in Table 2.

Due to the serious ID conversion problem in the tracking process, the improved algorithm with context information is compared with the original algorithm as shown in Table 3 below. The results show that after adding context information, the ID conversion rate in the video test decreases and achieves the expected effect.

| video | duration/s | number | density |
|----------|------------|--------|---------|
| sequence | | | |
| 1 | 63 | 19 | sparse |
| 2 | 10 | 15 | dense |
| 3 | 30 | 42 | dense |
| | | | |

Table 2. Traffic statistics experiment test video information.

| | headcount | ID | ID |
|----------------|-----------|------------|------------|
| video sequence | | conversion | conversion |
| | | (before) | (after) |
| 1 | 19 | 6 | 3 |
| 2 | 15 | 3 | 0 |
| 3 | 42 | 14 | 10 |
| | | | |

Table 3. Video test ID conversion comparison.

3.2.3 Experimental analysis of traffic statistics : After the target detection and target tracking tests are completed, we also tests the results of human flow detection. As shown in Table 4 below, we counted the number of positive detection, missed detection and false detection in each video, and calculated the recall rate, accuracy rate and F value of the video according to numerical values. As can be seen from the results, although the missed detection rate of pedestrians in dense places is high, the false detection rate is low. In general, the F value of the proposed method can reach more than 90%, showing a good performance.

| video sequence | Recall/% | Precision/% | F value/% |
|----------------|----------|-------------|-----------|
| 1 | 90.48 | 95 | 92.68 |
| 2 | 93.75 | 100 | 96.77 |
| 3 | 85.71 | 95.45 | 90.32 |

Table 4. Traffic video test results.

3.2.4 Experimental analysis of trajectory error correction : Based on the data collected by the experimental equipment, positioning experiments were conducted respectively according to the design scheme. In order to compare the accuracy of the positioning scheme , the real trajectory, P D R trajectory and



Figure 8. Error trajectory diagram.

modified trajectory were represented on the same graph, as shown in FIG. 8.

As can be seen from the trajectory result figure, the trajectory in the figure is similar to the original roadmap. After the error accumulates to a certain value, the real trajectory of PDR trajectory is larger, and finally fails to reach the closure of starting point and ending point. After adding coordinate correction, it can return to the original trajectory and further improve the accuracy of indoor pedestrian positioning.

4. CONCLUSION

This paper studies the pedestrian dead reckoning localization algorithm under the vision sensor, combined with the target detection and target tracking algorithm in computer vision and the inverse perspective transformation algorithm to correct the target track error. Experimental results show that the proposed algorithm can realize intelligent and efficient pedestrian detection on the basis of real-time detection, and correct errors in the process of pedestrian navigation prediction.

REFERENCES

Bewley A , Ge Z , Ott L , et al. Simple Online and Realtime Tracking [C]// 2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016.

Bodla N , Singh B , Chellappa R , et al. Improving Object Detection With One Line of Code[J]. 2017.

BOCHKOVSKIY A, WANG C Y,LIAO H Y M.YOLOv4; op timal speed and accuracy of object detection[EB/OL].[2020-04-23].https://github.com/AlexeyAB/darknet.

GIRSHICK R,DONAHUE J,DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision & Paltern Recognition, 2014.

Klein I , Solaz Y , Ohayon G . Pedestrian Dead Reckoning with Smartphone Mode Recognition[J]. IEEE Sensors Journal, 2018:1-1.

Liu W , Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector[J]. Springer, Cham, 2016.

Matthias Mueller, Neil Smith, Bernard Ghanem. Context-aware correlation filter tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR), IEEE ,2017 :1387-139 5.

Neubeck, A. and L. Gool. "Efficient Non-Maximum Suppression." 18th International Conference on Pattern Recognition (ICPR'06) 3 (2006): 850-855.

ROSS G.Fast R-CNN [J]. Computer Science,2015,9(1): 10-19.

Redmon J, Divvala S, Girshick R, et al. You only look once:Un ified, real-time object detection. Proceedings of the IEEEConfe rence on Computer Vision and Pattern Recognition.Vega s, NV, USA. 2016. 779–788.

Redmon J, Farhadi A. YOLO9000: better, faster, stronger. Proceedings of the IEEE Conference on Computer Visiona nd Pattern Recognition. Honolulu, HI, USA. 2017.7263– 7271. Redmon J, Farhadi A. Yolov3: An incremental improvement.ar Xiv preprint arXiv: 1804.02767, 2018.

REN S,HE K,GIRSHICK R,et al.Faster R-CNN: Towards realtime object detection with region proposal networks[J].IEEETransactions on Pattern Analysis & Machine Intelligence, 2017,39(6) : 1137-1149.

Wojke N , Bewley A , Paulus D . Simple Online and Realtime Tracking with a Deep Association Metric[J]. IEEE, 2017:3645-3649.

Yong Chen, Hao Wang, Yaqin Zhu, Haoliang Jia, Wei Wu. Multi-target tracking algorithm based on improved Deep Sort algorithm [J]. Application Research of Computers,2020,37(S2):311-312+315.

You-chun XU. Research on init ia tive security of vehicle based on machine Vision[D]. Beijing: Tsinghua University , 2003.

Yuqin, Wang, Peng, et al. Pedestrian Dead Reckoning-Assisted Visual Inertial Odometry Integrity Monitoring.[J]. Sensors (Basel, Switzerland), 2019, 19(24).

Yu Cao, Ying Feng, Bing Lei, Yuntao Yang, Lishuang Zhao. Error Analysis and Accuracy Verification of inverse Perspective Mapping Formula [J]. Acta Photonica Sinica,2011,40(12):1833-1838.

Zhang G, Sun X, Ren J, et al. Research on improved indoor positioning algorithm based on WiFi–pedestrian dead reckoning[J]. International Journal of Distributed Sensor Networks, 2019, 15(5):155014771985193.