Research Article

Research on Multitarget Detection and Intelligent Tracking Technology Based on Computer Vision

Xiaobo Liu [ID],[1] Junya Yan [ID],[1] and Yuxin Zhang[2]

[1]*Shanxi Vocational University of Engineering Science and Technology, Shanxi, Jinzhong 030619, China*
[2]*Taiyuan Institute of Technology, Shanxi, Taiyuan 030008, China*

Correspondence should be addressed to Xiaobo Liu; xiaobo.liu@sxgkd.edu.cn

With image analysis as the core for multitarget detection and intelligent tracking, mostly applying the Faster R-CNN or YOLO framework, the MOTA score for multitarget tracking is low in the face of complex working environments. Therefore, further research into computer vision techniques is carried out to design new multitarget detection and intelligent tracking methods. Based on the small-aperture imaging model, the principle of lens distortion was analyzed, and a camera calibration and image calibration scheme was designed to obtain effective environmental images. The attention mechanism is introduced to optimise the structure of deep learning networks, and a computer vision detection algorithm based on this is applied to complete regional multitarget detection. The distance between each target and the body is then measured in combination with binocular vision principles. Finally, the spatiotemporal context algorithm is applied to perform simulation calculations to obtain the multitarget intelligent tracking results. The experimental results show that the mean MOTA score of the proposed technique is 0.87 in the night environment, which is 24.14% and 28.374% better than the neural network-based and machine vision-based tracking methods, respectively; in the daytime environment, the mean MOTA score of the multitarget tracking results of the technique is 0.94, which is 28.72%, and the mean MOTA score of 0.94 for the multitarget tracking results in the daytime environment was 28.72% and 22.34% higher than the other two methods.

1. Introduction

The progress of modern technology has led the automotive industry towards an intelligent trend [1], and it can be said that the level of development of multitarget detection and intelligent tracking technology directly determines the degree of intelligence of the vehicle. As urban traffic scenarios become more and more complex [2], the technical requirements for intelligent driving are becoming higher and higher. To ensure the stability and safety of automated vehicle driving, it is necessary to first detect targets in the driving area and track their movement trends in order to generate highly intelligent driving decisions [3]. At the same time, a key aspect of smart driving vehicle operation is environmental perception. Only a clear enough knowledge of the road environment around the driving area can ensure

that smart driving vehicles are integrated into the traffic environment. However, existing multitarget detection and intelligent tracking technologies can be affected by good or bad lighting and weather conditions, resulting in target detection and intelligent tracking effects that do not meet intelligent driving requirements.

With the in-depth research of computer vision technology, computer vision technology with deep learning as the core has started to be applied in various fields. The paper takes this as the research direction and introduces an attention mechanism to further optimise the current computer vision technology by adding a subregion feature library and an aspect ratio feature library to the original detection model to improve the feature representation capability of the computer vision-based multitarget detection model and obtain accurate localisation recognition results.

On this basis, a new multitarget detection and intelligent tracking technique is established to stably and quickly track the movement trend of multitarget objects in the vehicle's surrounding environment. The experimental validation results show that the application of the proposed technique for multitarget tracking results in more accurate tracking.

This paper consists of four chapters. The first chapter is the introduction. The second chapter introduces the design of multitarget detection and intelligent tracking based on computer vision. The third chapter is the empirical analysis, and the designed multi-target detection and intelligent tracking algorithm is tested. The fourth chapter is the conclusion.

2. Multitarget Detection and Intelligent Tracking Technology

2.1. Camera Calibration and Image Calibration Programme. The first step in tracking is identification. A binocular camera with a wide-angle lens is mounted on the front of the smart driving vehicle as the main device for sensing the smart driving environment. Considering that radial and tangential aberrations exist when the camera captures images [4], a camera calibration and image calibration scheme is established with the objective of reducing the position errors caused by lens aberrations as a fundamental part of computer vision-based multitarget detection and intelligent tracking [5]. Most of the imaging of the camera relies on the small-aperture imaging model, in three-dimensional space and the camera imaging plane, respectively, to determine a target point with a corresponding relationship between the two can be expressed as follows:

$$I[a'\,b'\,1] = [a\,b\,c\,1]\begin{bmatrix} R \\ T \end{bmatrix}, \tag{1}$$

$$[g\,h\,1] = [a'\,b'\,1]\psi. \tag{2}$$

In the formula, (a, b, c) represents the coordinates of the target point in the world coordinate system, (a', b') represents the coordinates of the target point in the camera coordinate system, (g, h) represents the coordinates of the target point in the image pixel coordinate system, I represents the scaling factor, R represents the rotation matrix, T represents the translation matrix, and ψ represents the internal parameter matrix.

$$\psi = \begin{bmatrix} d_1 & 0 & 0 \\ \tau & d_2 & 0 \\ u_0 & v_0 & 1 \end{bmatrix}. \tag{3}$$

In the formula, d_1 and d_2 represent the scale factor on the x and u axes, respectively, (u_0, v_0) represents the coordinates of the origin, and τ represents the axis inclination parameter. With (1) and (2), the imaging principle of the camera is described directly, but in practice, a nonlinear aberration model needs to be added to describe the imaging point shift.

$$\begin{bmatrix} \overline{\varepsilon} \\ \overline{o} \end{bmatrix} = \left(1 + \varphi_1 r^2 + \varphi_2 r^4 + \varphi_3 r^6\right)\begin{bmatrix} \overline{\varepsilon_1} \\ \overline{o_1} \end{bmatrix}$$
$$+ \begin{bmatrix} 2k_1\overline{\varepsilon_1 o_1} + k_2\left(r^2 + 2\overline{\varepsilon_1}^2\right) \\ 2k_2\overline{\varepsilon_1 o_1} + k_1\left(r^2 + 2\overline{o_1}^2\right) \end{bmatrix}. \tag{4}$$

In the equation, $(\overline{\varepsilon}, \overline{o})$ represents the ideal coordinate value of the target point, $(\overline{\varepsilon_1}, \overline{o_1})$ represents the actual distortion coordinate value of the target point, φ_1, φ_2, and φ_3 represents the radial distortion factor, r represents the radius, and k_1 and k_2 represents the tangential distortion factor. Based on the exact and distorted coordinates, the camera calibration process is completed and reasonable parameter values are obtained.

In practice, it is necessary to combine the Zhang Zhengyou calibration method with a calibration plate composed of black and white squares to take images, build a library of photos for calibration, apply the optimization method to iteratively solve the camera parameters with the minimum error as the goal [6], determine the optimal internal and external parameters and aberration parameters of the camera, and after the camera calibration is completed, calculate the aberration pixel coordinates of the actual photos taken, as well as the ideal coordinates to complete the image calibration.

2.2. Computer Vision Multitarget Detection Algorithms. A calibrated camera is applied to capture images of the smart driving surroundings, and then computer vision techniques are applied for multitarget detection. Considering the computer vision detection technology based on deep learning, it can be influenced by external factors in practical applications, making the detection results biased. The paper applies computer vision principles and introduces attention mechanisms in conventional deep learning networks [7] to establish the multitarget detection framework shown in Figure 1 for feature extraction, feature pooling, and classification regression of camera acquisition images.

According to Figure 1, it can be seen that applying convolutional neural network for computer vision multi-target detection requires first acquiring image convolutional features for classification and regression analysis. In the paper, the subregion feature attention module and aspect ratio feature attention module [8] are introduced into the original multitarget detection model to obtain the computer vision target detection block diagram shown in Figure 2. Introducing subregion feature attention module and aspect ratio feature attention module into the original multitarget detection model can help distinguish single target in multitarget better. Therefore, higher recognition accuracy can be obtained for target recognition.

According to Figure 2, the updated computer vision target detection model contains the attention module, which mainly plays a role in the ROI feature extraction process, extracting features for regular patterns for further processing and combining them with the original ROI pooled features to generate high-quality ROI classification features.
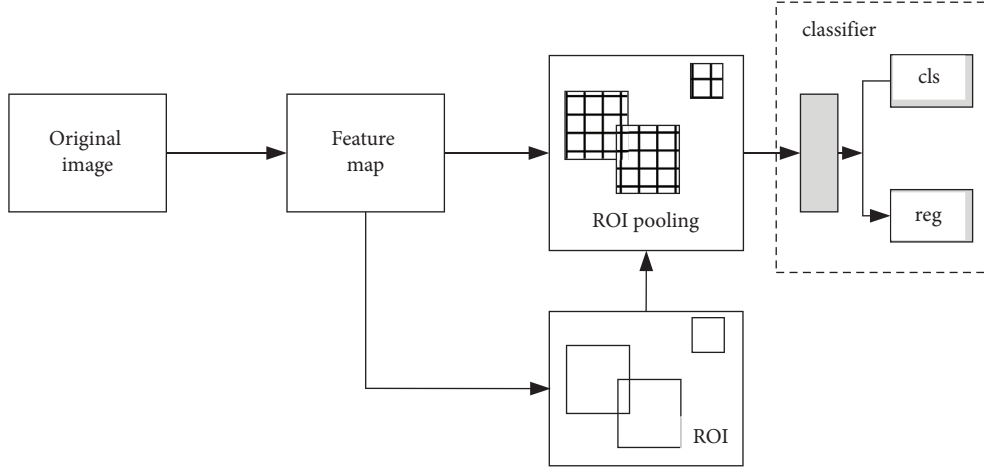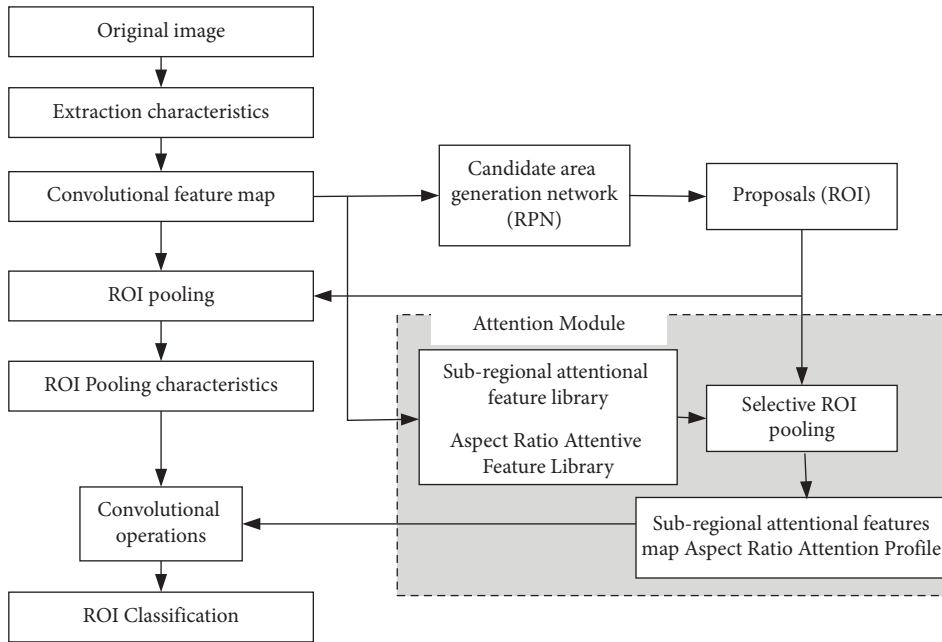
Figure 1: Block diagram of multitarget detection.

Figure 2: Block diagram of computer vision target detection with the introduction of an attention module.

$$\xi B_p = B_p \bullet (M_1 + M_2). \tag{5}$$

In the formula, p represents ROI, B_p represents ROI pooling features, ξ represents ROI classification features, M_1 represents subregion attention feature maps, and M_2 represents aspect ratio attention feature maps.

The new attention module contains two attention feature bases, each of which holds the corresponding attention feature activation relationships. Among them, the features displayed within the subregion attention feature library are associated with spatial location information [9], and based on the location of each feature point in the ROI subregion, the formula for calculating the subregion attention salience value is expressed as follows:

$$\mu_\beta (i, j) = \phi U_{ij}. \tag{6}$$

In the formula, (i, j) denotes a point in the convolutional feature map, β denotes a subregion, μ denotes an attentional salient, U denotes a feature vector, and ϕ denotes an attentional feature extractor.

The features saved within the aspect ratio attention feature library describe actual feature attributes that directly describe the observed viewpoint and pose morphology of the target object, extracting horizontal and vertical scale differences in the target detection framework and better determining target class differences. A deep learning network structure incorporating attention mechanisms is applied to run computer vision techniques to obtain multitarget detection results during intelligent driving.
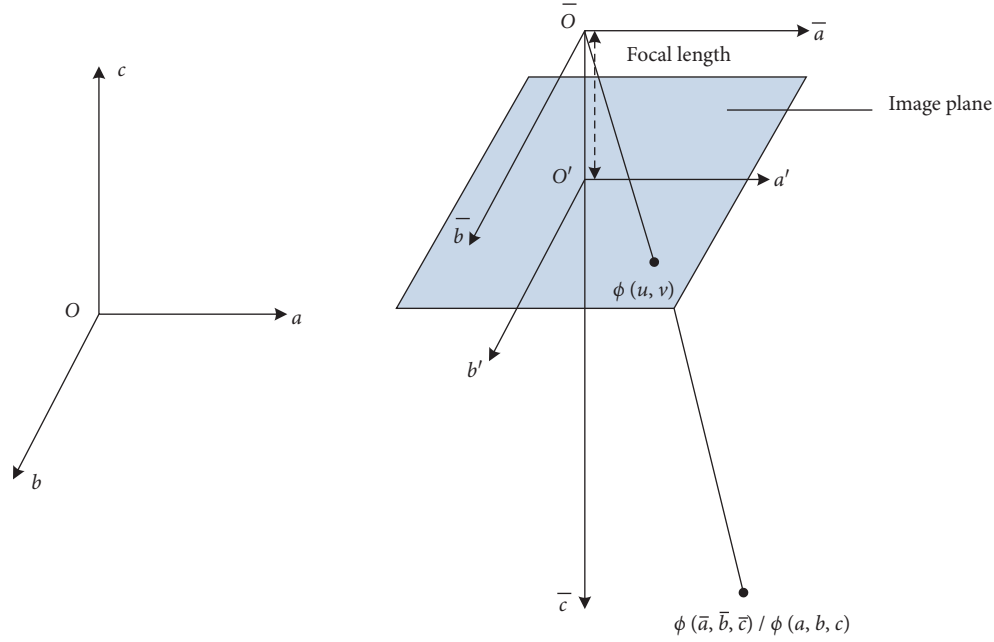
FIGURE 3: Diagram of each coordinate system of the pinhole model.

2.3. Target Distance Measurement Programme. After the target detection is over, the target distance measurement method is designed based on the binocular vision principle to locate the distance between each target point and the vehicle body. Applying the pinhole model imaging principle [10], image acquisition is carried out during intelligent driving, and each coordinate system of the pinhole model is shown in Figure 3.

In Figure 3, $(Oabc)$ represents the 3D world coordinate system, $(\overline{O}\overline{a}\overline{b}\overline{c})$ represents the camera coordinate system, $(O'a'b')$ represents the image plane coordinate system, $\phi(u,v)$ represents the plane coordinate points, and $\phi(\overline{a},\overline{b},\overline{c})\phi(a,b,c)$ represents the 3D coordinate points.

Setting up the existence of a target point in space and the known coordinates of the image point of the target point in the two calibrated camera coordinate systems, combined with the projection matrix, the perspective projection matrix transformation relationship can be expressed as follows:

$$\overline{c}_A \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} = \chi_1 \begin{bmatrix} \overline{a}_L \\ \overline{b}_L \\ \overline{c}_L \\ 1 \end{bmatrix}, \tag{7}$$

$$\overline{c}_A \begin{bmatrix} u_A \\ v_A \\ 1 \end{bmatrix} = \chi_2 \begin{bmatrix} \overline{a}_A \\ \overline{b}_A \\ \overline{c}_A \\ 1 \end{bmatrix}. \tag{8}$$

In the formula, L denotes the left camera, A denotes the right camera, χ_1 denotes the projection matrix of the left camera, and χ_2 denotes the projection matrix of the right camera. Based on the left and right camera perspective projection matrix transformation relationships shown in (7) and (8), the spatial coordinates of the target can be deduced from the known image point coordinates and the distance measurement results can be obtained by comparing the coordinate information. It is important to note that camera images in complex environments can contain a lot of noise, which can affect the accuracy of the distance measurement results. In this case, the least squares method can be combined with further solutions to obtain more accurate spatial coordinates of the measured point.

2.4. Multiobjective Intelligent Tracking. Relying on computer vision technology, the correlation between the target object and the local scene needs to be analysed for intelligent tracking of multiple targets for detection and localisation [11], and the spatiotemporal context algorithm based on Bayesian framework is applied in the paper for simulation and calculation to clarify the intensity and location correlation of the target region in the image of the local context, and then the maximisation confidence function [12] is applied to achieve the target location of real-time tracking. In which, the confidence function can be expressed as follows:

$$e(X) = \Gamma \times E^{-\left(\frac{|X - X''|}{\partial}\right)^{\Theta}}. \tag{9}$$

In the formula, X is the target position, X'' is the target region centre position, e is the confidence level, Γ is the normalisation factor, E is the bias function, ∂ is the scale parameter, and Θ is the shape parameter.

Faced with a multiframe image acquired in real time, the local contextual feature set of the target region it contains can be represented as follows:
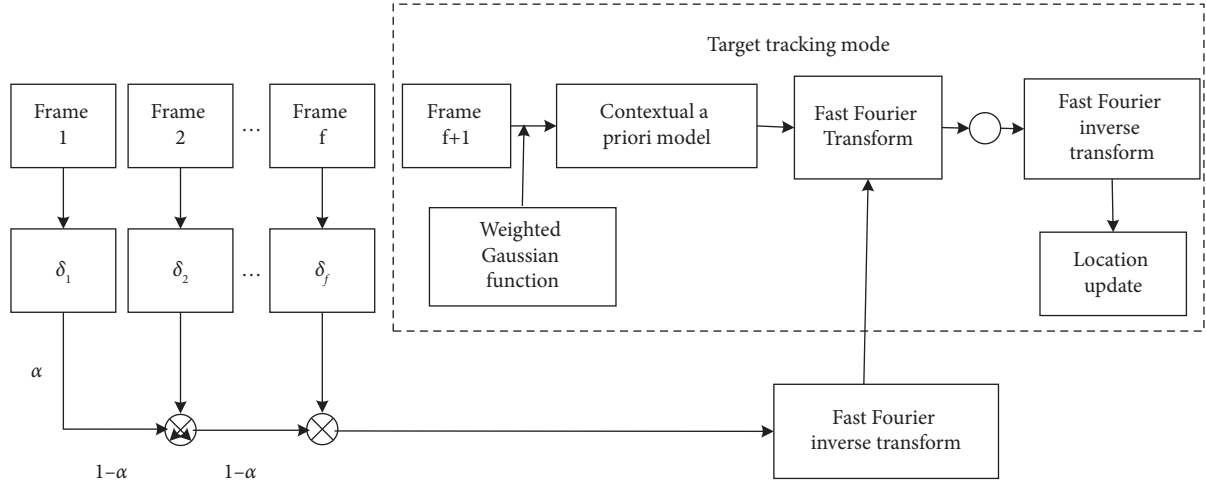
FIGURE 4: Multitarget intelligent tracking model.

$$D = \{e(\Upsilon) - G(\Upsilon), \Upsilon \in \Omega\}. \qquad (10)$$

In the formula, D represents the local context feature set, Υ represents the selected location, G represents the image intensity, and Ω represents the local context region around the selected location. Based on the concept of local context analysis [13], the multiobjective intelligent tracking model can be described as Figure 4.

According to Figure 4, the STC algorithm-based multitarget intelligent tracking model is essentially a multitarget intelligent tracking task through a maximum confidence function search. The target image frame is analysed to obtain its corresponding spatial context model, and the spatio-temporal context model of the next image frame is represented as follows:

$$H_{f+1}(X) = (1 - \alpha)H_f(X) + \alpha\delta_f(X). \qquad (11)$$

In the formula, f denotes the number of image frames, H denotes the spatiotemporal context model, α denotes the learning rate factor, and δ denotes the spatial context model. According to (11), it can be seen that the spatiotemporal context model of the next frame can be derived by weighting the spatiotemporal context model for the current frame of the photo and the spatial context model [14]. Therefore, the confidence function calculation formula can be updated as follows:

$$e_{f+1}(X) = Q^{-1}\big(Q\big(H_{f+1}(X)\big) \otimes Q\big(G_{f+1}(Y)w_\partial(X - X'')\big)\big). \qquad (12)$$

In the formula, Q denotes the Fourier transform, Q^{-1} denotes the Fourier inverse transform, $\otimes$ denotes the convolution operation, and w denotes the weighted Gaussian function. Based on the results of the maximum confidence calculation, the position of the target point within each frame is determined. Considering the application of the STC algorithm, it is only possible to describe the change in the target position of a single pixel and obtain the actual displacement of the pixel. To obtain more intuitive target tracking results, it is also necessary to fuse optical flow algorithms [15] to specify the subpixel displacement in consecutive image frames and obtain multitarget intelligent tracking results.

3. Experiment

The design of a multitarget detection and tracking technique is based on computer vision technology, and an experimental analysis is carried out to verify the effectiveness of the technique in practice. A binocular camera is mounted on an ordinary car to perform multitarget detection and tracking in night and day scenes. Based on the detection and tracking results, the validity of the research content in the paper is reflected.

3.1. Camera Calibration.
The implementation of computer vision technology needs to be image-based. During the experiments, a camera calibration process is carried out before the camera is fixed to the car in order to capture a more realistic image. First, create a $13 * 14$ black and white checkerboard visual calibration grid, each measuring $20\,mm * 20\,mm$, as shown in Figure 5. Print out the calibration grid and paste it onto a flat horizontal board to form a visual calibration board.

When applying the binocular camera to capture images, it is necessary to constantly adjust the angle of the visual calibration plate to obtain multiple calibration plate images, as shown in Figure 6.

The calibration plate images captured by the cameras shown in Figure 6 were loaded simultaneously into the MATLAB software and manually processed through the Calibration Toolbox to extract the corner points contained within each calibration plate image, and the optimum internal and external parameters for the left and right cameras were calculated based on the image corner point information, as shown in Tables 1 and 2.

Among them, η_x, η_y indicates the camera focal length parameter, (u_0, v_0) indicates the camera principal point
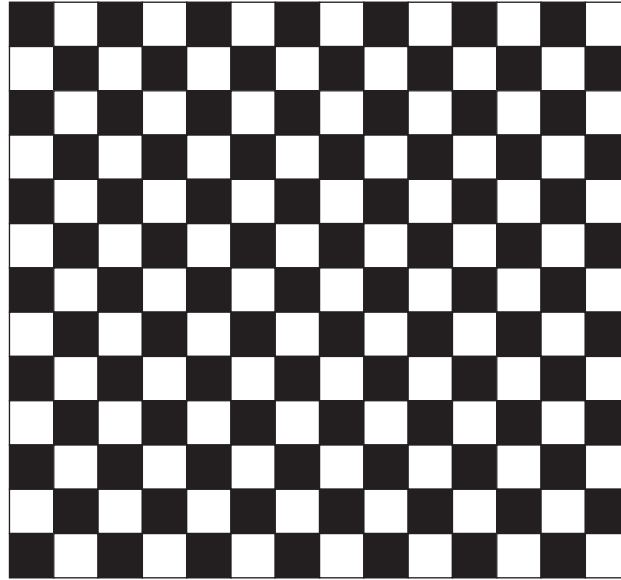
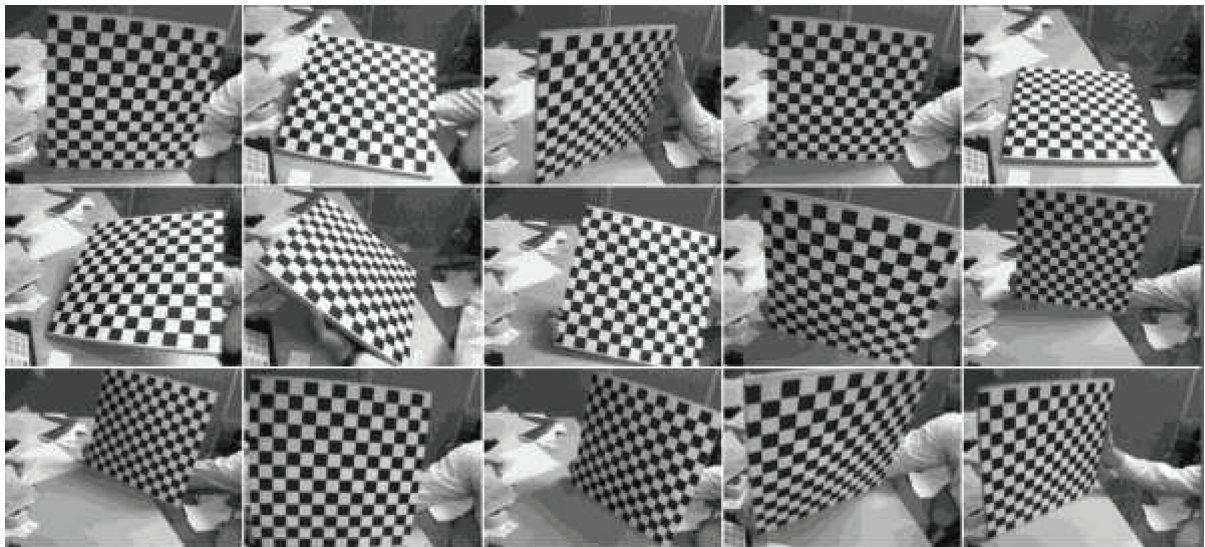FIGURE 5: Schematic diagram of the visual calibration grid plate.

FIGURE 6: Image information of the calibration plate captured by the camera.

TABLE 1: Statistics of in-camera parameters.

Internal parameter	Left camera calibration result	Right camera calibration result
η_x	5523.45	5528.86
η_y	5523.57	5512.53
u_0	775.98	774.43
v_0	631.52	543.78
l_1	−0.137	−0.162
l_2	3.454	0.338
s_1	0.006	−0.001
s_2	−0.001	0.008

TABLE 2: Statistics of off-camera parameters.

External parameter	Calibration result
x-Axis rotation matrix (Rx)	$[0.98\ 0.07\ 0.03]^{\mathrm{T}}$
y-Axis rotation matrix (Ry)	$[-0.07\ 0.98\ -0.02]^{\mathrm{T}}$
z-Axis rotation matrix (Rz)	$[-0.03\ 0.02\ 0.98]^{\mathrm{T}}$
Translational variables (T)	$[-303.02\ -0.54\ 6.63]$

position, (l_1, l_2) indicates the radial aberration, and (s_1, s_2) indicates the tangential aberration. Through the camera calibration process, in addition to the intracamera parameters shown in Table 1, the external parameters of the right

camera relative to the left camera were also obtained, as shown in Table 2.

After the parameters inside and outside the camera have been adjusted, it is fixed to the vehicle and continuous image acquisition is carried out while the vehicle is in motion as the data required for multitarget detection and intelligent tracking.

FIGURE 7: Night multitarget tracking results.

3.2. Evaluation Indicators.

In order to demonstrate the reliability of the design technique in the paper, the Multi-Objective Tracking Accuracy (MOTA) should be selected during this experiment to assess the consistency between the intelligent tracking results and the actual trajectory of the target. The intelligent tracking result output from the application of the techniques in the text is first obtained to form a tracking path containing multiple nodes, and then the true position of each target is investigated to generate an actual running path containing multiple nodes. Comparing the degree of matching between the two paths and analysing the false detections, misclassifications, and incorrect matches that occur during the tracking process, the MOTA score is calculated by expressing the formula as follows:

$$MOTA = 1 - \frac{(m + n + \gamma)}{N}. \tag{13}$$

In the formula, m represents the number of false targets detected during tracking, n represents the number of false targets, γ represents the number of false match targets, and N represents the number of all targets that appear in the image frame.

3.3. Visual Analysis of Tracking Results.

Firstly, a complex scene with high footfall was selected as the experimental scene. The experimental vehicle is controlled to drive through the scene and take video in a night environment, and several images are captured within the video sequence to form a multitarget tracking dataset. Using the multitarget detection and intelligent tracking techniques proposed in the paper, the dataset was analysed to obtain the multitarget tracking results shown in Figure 7.

Figure 7 shows frames 3, 10, 15, and 25, from which it can be seen that the scene contains a large number of pedestrians and that the pedestrian trajectories are not identical, while the tracking technique proposed in the text, when applied, detects essentially all the pedestrians, except in cases where the occlusion is too severe. The comparison of the four frames shows that the position of the pedestrian changes within each frame, but the colour of the detection frame does not change, which indicates a better result for multitarget tracking. The above experiments show that the tracking technique designed in the paper can detect and track multiple targets accurately in complex scenarios.

Afterwards, a weather day with good lighting conditions was selected to conduct a multitarget detection and intelligent tracking experiment on a relatively secluded street. The vehicle is set to drive through the street at an even speed and the binocular camera captures images of the surrounding scene as it moves to form a second experimental dataset. This dataset is processed by applying the techniques in this paper to obtain the multitarget tracking results shown in Figure 8.

As can be seen from Figure 8, there are only a small number of vehicles moving in the area and no pedestrians passing by. Therefore, the car can be considered as a detection target, and the illustration captures frames 18, 26, 32, and 40 to visualise the tracking results of a moving vehicle. Overall, the proposed technique can be applied to quickly detect other vehicles from the moment they come into the camera's range and mark them with different coloured detection boxes and then keep track of the vehicle's movement until it leaves the camera's range. In addition, although the experiment was conducted during daylight hours with few targets, the lighting conditions were complex, containing both shaded and brightly lit areas. However, the detection and tracking results of the techniques designed in the paper were not affected by the light and showed that the application of computer vision techniques based on deep learning networks with the introduction of attention mechanisms ensures the reliability of multitarget detection and tracking.

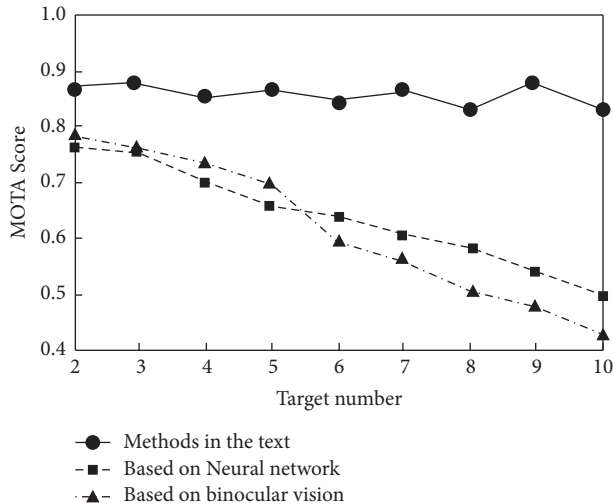FIGURE 8: Daytime multitarget tracking results.

FIGURE 9: Comparison of MOTA scores for different methods in the night environment.

FIGURE 10: Comparison of MOTA scores for different methods in a daytime environment.

3.4. Tracking Performance Comparison. To enhance the visualisation of the experimental results, simultaneous experimental analyses were carried out in different experimental scenarios by applying the techniques mentioned in the text, a neural network-based approach, and a binocular vision-based approach, respectively. The targets to be detected and tracked were set to keep increasing, and the detection and tracking results of the different methods were recorded. The variation of MOTA scores for different methods in the night environment is shown in Figure 9.

According to Figure 9, the MOTA scores of the tracking results of the designed methods in the paper do not fluctuate much after the number of targets increases, with an average value of 0.87. The MOTA scores of the other two methods, however, keep decreasing as the number of targets increases. The neural network-based approach reduced the MOTA score from 0.76 to 0.52, with an average MOTA score of 0.66, while the binocular vision-based approach for multitarget
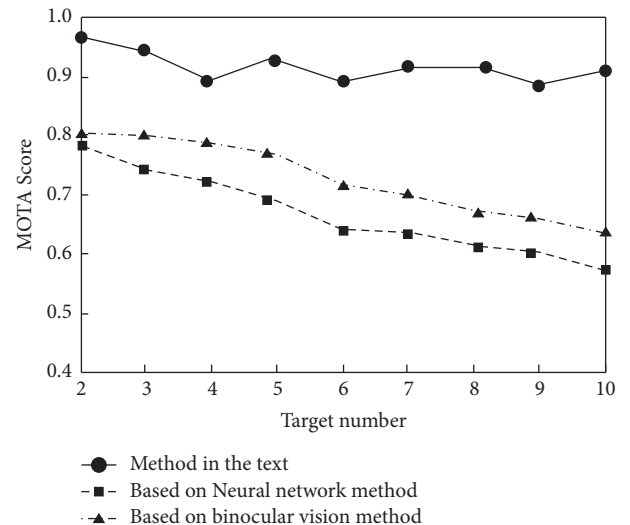
detection and tracking in different environments achieved a maximum MOTA score of 0.78 and a minimum of 0.43, with an average MOTA score of 0.62. In summary, the MOTA scores of the intelligent tracking results of the proposed technique improved by 24.14% and 28.374% compared to the other two methods.

The change in MOTA scores for the different methods in the daytime environment was then analysed to form the comparative results of MOTA scores shown in Figure 10.

According to Figure 10, the mean MOTA score for the multitarget tracking results of the proposed technique in the daytime environment is 0.94, while the mean MOTA scores of the other two methods are 0.67 and 0.73, respectively. As a result, the proposed technique improves MOTA scores by 28.72% and 22.34% compared to neural network-based and binocular vision-based methods, respectively.

4. Conclusion

Intelligent driving is the development trend of the future automobile industry. The application of intelligent driving in urban scene depends on the development level of dynamic object tracking technology.

In this paper, the traditional computer vision technology based on deep learning is optimized. By adding attention mechanism, this paper constructs a new multiobject detection algorithm in computer vision to achieve more accurate and fast multiobject detection. Moreover, using the STC model, the results of the intelligent target tracking algorithm established in this paper are also more accurate.

The intelligent tracking method designed in this paper can maintain accurate detection and tracking of multiple targets in complex traffic environment. It performs better than the traditional intelligent tracking method in the environment of multitarget and changing illumination conditions. The intelligent tracking method constructed in this paper is applied to the field of intelligent driving, which is beneficial to enhance the stability of vehicle driving.

Data Availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] X. Wang, W. Xie, and L. Li, "Labeled multi-Bernoulli maneuvering target tracking algorithm via TSK iterative regression model," *Chinese Journal of Electronics*, vol. 31, no. 2, pp. 227–239, 2022.

[2] Y. Zhu, W. Zhou, Q. Li, and L. Yu, "Parallax-estimation-enhanced network with interweave consistency feature fusion for binocular salient object detection," *IEEE Signal Processing Letters*, vol. 28, pp. 927–931, 2021.

[3] Y. Su, F. Tao, J. Jin, and C. Zhang, "Automated overheated region object detection of photovoltaic module with thermography image," *IEEE Journal of Photovoltaics*, vol. 11, pp. 535–544, 2021.

[4] J. Li, J. Su, C. Xia, M. Ma, and Y. Tian, "Salient object detection with purificatory mechanism and structural similarity loss," *IEEE Transactions on Image Processing*, vol. 30, pp. 6855–6868, 2021.

[5] H. Wang, S. Liao, and L. Shao, "AFAN: augmented feature alignment network for cross-domain object detection," *IEEE Transactions on Image Processing*, vol. 30, pp. 4046–4056, 2021.

[6] W. Yi, Z. Fang, W. Li, R. Hoseinnezhad, and L. Kong, "Multiframe track-before-detect algorithm for maneuvering target tracking," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4104–4118, 2020.

[7] S. T. Moon, W. Youn, and H. Bang, "Novel deep-learning-aided multimodal target tracking," *IEEE Sensors Journal*, vol. 21, pp. 20730–20739, 2021.

[8] B. Tuncer and E. Ozkan, "Random matrix based extended target tracking with orientation: a new model and inference," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1910–1923, 2021.

[9] R. B. Angle, R. L. Streit, and M. Efe, "Multiple target tracking with unresolved measurements," *IEEE Signal Processing Letters*, vol. 28, pp. 319–323, 2021.

[10] A. K. Roonizi, "An efficient algorithm for maneuvering target tracking [tips & tricks]," *IEEE Signal Processing Magazine*, vol. 38, no. 1, pp. 122–130, 2021.

[11] D. Mery, D. Saavedra, and M. Prasad, "X-ray baggage inspection with computer vision: a survey," *IEEE Access*, p. 1, 2020.

[12] M. K. I. Lee, M. Rabindranath, K. Faust et al., "Compound computer vision workflow for efficient and automated immunohistochemical analysis of whole slide images," *Journal of Clinical Pathology*, vol. 5, no. N, pp. 9–2021-208020, 2022.

[13] S. Buzzi, E. Grossi, M. Lops, and L. Venturino, "Radar target detection aided by reconfigurable intelligent surfaces," *IEEE Signal Processing Letters*, vol. 28, pp. 1315–1319, 2021.

[14] H. Lian, X. Pei, and X. Guo, "A local environment model based on multi-sensor perception for intelligent vehicles," *IEEE Sensors Journal*, vol. 21, pp. 15427–15436, 2021.

[15] Y. Zhang, Z. Lu, D. Ma, J. H. Xue, and Q. Liao, "Ripple-GAN: lane line detection with ripple lane line detection network and wasserstein gan," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1532–1542, 2021.