REAL-TIME REMOTE MONITORING AND TRACKING TECHNOLOGY FOR THREE-DIMENSIONAL WAREHOUSE BASED ON DEEP LEARNING AND TARGET DETECTION

Shaokui YAN[1], Haili DING[2,*], Jie ZHANG[3], Xiangwei WANG[4], Mingqiang WANG[5]

In order to realize the dynamic management of goods in the warehouse and improve the overall management efficiency and safety performance of the warehouse, a deep learning and target detection technology for remote real-time monitoring and tracking of the stereo warehouse is proposed. Wireless sensor nodes are used to provide all-round coverage of the three-dimensional warehouse to meet the monitoring needs of various parameters such as cargo status and cargo location, and the image data of the three-dimensional warehouse is trained by deep confidence networks to extract high-quality and distinguished features of the three-dimensional warehouse. The hybrid Gaussian model is used to accurately locate and identify the target warehouse, determine the location and trajectory of warehouse objects, and finally introduce a priori and measurement information by virtue of Kalman filtering method to realize real-time monitoring and tracking of the stereo warehouse. The results show that the AUC value of the complex monitoring attributes of the proposed method is as high as 0.93 after the optimization process, and the monitoring accuracy is as high as 98.43%, and the monitoring time is short, and the predicted value of the center of mass coordinates of the warehouse is the same as the real value of the actual coordinates, which indicates that the proposed method can ensure the comprehensiveness and real-time monitoring and tracking, and realize the intelligence and efficiency of the warehouse management.

Keywords: deep confidence network; hybrid Gaussian model; target detection; three-dimensional warehouse management; real-time monitoring; tracking and prediction

1. Introduction

As an efficient storage and management method, stereo warehouses are widely used in logistics and warehouse management [1-3]. In today's digital and

[1] Eng., State Grid Ningxia Marketing Service Center (State Grid Ningxia Metrology Center), Yinchuan, China, E-mail: yanshaokui@yxzx.nx.sgcc.com.cn

[2*] Eng., State Grid Ningxia Marketing Service Center (State Grid Ningxia Metrology Center), Yinchuan, China, Corresponding e-mail: dinghaili@yxzx.nx.sgcc.com.cn

[3] Eng., State Grid Ningxia Marketing Service Center (State Grid Ningxia Metrology Center), Yinchuan, China, E-mail: zhangjie@yxzx.nx.sgcc.com.cn

[4] Eng., Comarvel Intelligent Technology Company Limited, Yantai, China, E-mail: 286595263@qq.com

[5] Eng., Comarvel Intelligent Technology Company Limited, Yantai, China, E-mail: 547606201@qq.com

intelligent mega-trend, warehouse management faces more efficient and accurate requirements, especially for large stereo warehouses, the need for real-time monitoring and tracking technology is more urgent [4]. Warehouse monitoring is a measure taken to ensure the safety, efficiency and compliance of warehouses. Through warehouse monitoring, it can realize all-around real-time monitoring and management of goods, personnel, equipment, environment, etc. inside the warehouse, so as to improve the management efficiency of the warehouse, reduce costs, guarantee the safety of goods and comply with regulatory requirements. However, in the traditional three-dimensional warehouse, the warehouse management method still relies too much on manual operation, and it is unable to monitor the changes and abnormalities of the goods inside the warehouse in real time, and there are potential safety hazards, which leads to bottlenecks in the supply chain, logistics management and other aspects that cannot be effectively solved [5-7].Therefore, it is especially important and necessary to research and develop a remote real-time monitoring and tracking technology based on deep learning and target detection for three-dimensional warehouses. This technology can provide all-round coverage of the three-dimensional warehouse through wireless sensor nodes, meet the monitoring needs of various parameters such as cargo status and cargo location, realize the dynamic management of the warehouse cargo, and improve the overall management efficiency and safety performance of the warehouse.

For example, literature [8] proposed an urban water supply pollution source anti-tracking technique based on water quality monitoring, constructed an auxiliary graph theory model of the pipe network, calculated the shortest flow path between any two nodes, obtained the minimum number of covered nodes by greedy algorithm, and gave the optimal layout to complete monitoring and tracking. In the literature [9], the offset tracking method using SAR image intensity can detect large deformation and make up for the deficiency in large deformation object monitoring, and a deformation monitoring method based on deformation gradient and image noise is proposed, which considers the influence of deformation gradient and noise on deformation, and adaptive CCW is selected to perfectly monitor and track the deformed object objects. A tracking algorithm in the presence of severe IR interference was developed in the literature [10]. A cascaded R-CNN based detection method is used to detect the target and the situational assessment is performed based on the detection results. The location of the target is estimated from the detection results, and accurate monitoring and tracking of multiple targets is achieved. The literature [11] investigated an IGBT online monitoring method based on artificial neural networks. The practical measurable parameters DC link voltage and H-bridge output voltage were selected as the inputs of the neural network, and single-input-single-output and multiple-input-single-output neural networks were analyzed and discussed. Using this method, the practical measurable parameters can be established and then accurately monitored and tracked in real time for the target object. Literature [12] proposes super-resolution DOA algorithms using uniform linear arrays in the presence of white noise for target detection and tracking by estimating the direction and number of waves incident on the antenna elements over a specific frequency range. The concept of DOA estimation and its mathematical modeling for each method are also elucidated. MATLAB simulations are used to model each DOA computation method for various scenarios to evaluate its performance in order to obtain the required accuracy and resolution of each DOA algorithm. Literature [13] proposes an improved genetic optimization based indoor target tracking method, this resampling method optimizes the distribution of resampled particles through five operators, namely selection, roughening, classification, crossover and mutation. The proposed resampling method is then integrated into a particle filtering framework to form a particle filter based on genetically optimized resampling, which improves localization accuracy and robustness.

The traditional warehouse management method relies too much on manual operation, which cannot meet the requirements of modern warehouse management for efficiency, accuracy and real-time. At the same time, the monitoring and tracking of the status of items inside the warehouse also faces problems such as light changes and target blocking, which further increases the difficulty of real-time monitoring and tracking. This paper proposes a more accurate and robust remote real-time monitoring and tracking technology for three-dimensional storage based on deep learning and target detection algorithms to meet the practical needs of storage management. Firstly, we use wireless sensor nodes to cover the stereo warehouse in all directions to meet the requirements of monitoring various parameters such as cargo status and cargo location, train the image data of stereo warehouse through deep confidence network, and then extract high quality and differentiated features of stereo warehouse. Combined with the hybrid Gaussian model for accurate localization and identification of the target warehouse. The objects in the warehouse are detected and their positions and trajectories are determined. Finally, the Kalman filtering method is used to introduce a priori and measurement information to achieve real-time monitoring and tracking of the stereo warehouse.

2. Stereo warehouse rectangular node deployment pre-processing

Traditional warehouse environment monitoring systems suffer from numerous on-site devices, complex wiring, low reliability, and high management and maintenance costs. Therefore, covering the wireless sensor network in the three-dimensional warehouse can meet the monitoring needs of various parameters such as temperature, humidity, air, cargo status, and cargo location, which provides an excellent basis for the subsequent monitoring and tracking of the three-dimensional warehouse [12].

Set the sensing radius of the node of the wireless sensor as r and the cube with the side length of the depot as the monitoring area of the three-dimensional depot. With the center of the cube depot, point O, as the center centroid of the circle, the following constraints need to be satisfied by the perceived radius and a cube:

$$r \geq \sqrt{3} / 2a \tag{1}$$

The nodal sphere can cover the cube in all directions, indicating that the monitoring area of the stereo warehouse can be completely covered by wireless sensor nodes. Due to the shape of the stereo warehouse, the network nodes are deployed through a grid partitioning method. Set the length of the stereo warehouse as L, the width as W and the height as H. Partition the stereo warehouse into small cubes of a, i.e. $a \leq 2\sqrt{3} / 3r$.

The coverage expression of wireless node deployment in the cubic warehouse is shown in Fig. 1. The node is deployed in the center of the cube, which can complete the single-weight coverage of the whole cubic warehouse.

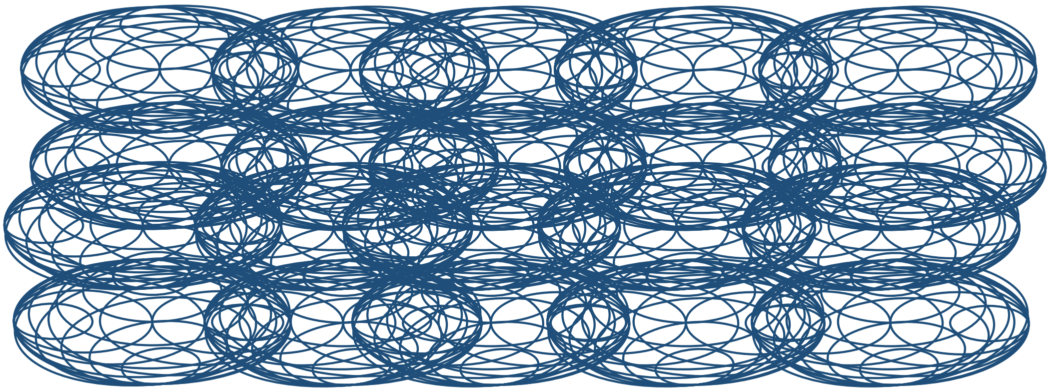


Fig. 1 Deployment coverage characterization of wireless nodes in the stereo warehouse

The expression for calculating the number of nodes of a wireless sensor is as follows.

$$N = \left\lceil \frac{L}{a} \right\rceil * \left\lceil \frac{W}{a} \right\rceil * \left\lceil \frac{H}{a} \right\rceil \tag{2}$$

The minimum coefficient N that needs to be solved to carry out the depot wireless node deployment, when $a = 2\sqrt{3} / 3r$ then the single heavy all-round coverage value N of the depot can be satisfied, expression:

$$\min N = \left\lceil \frac{L}{2\sqrt{3}/3r} \right\rceil * \left\lceil \frac{W}{2\sqrt{3}/3r} \right\rceil * \left\lceil \frac{H}{2\sqrt{3}/3r} \right\rceil \tag{3}$$

3. Remote real-time monitoring and tracking of stereo warehouses

3.1 Deep learning-based image feature extraction of stereo warehouse

First of all, a large number of warehouse monitoring images need to be collected, using Hikvision's DS-2CD3332-E model monitoring sensor, the camera has a high-definition image quality, intelligent analysis, wide dynamic, fog permeability, glare suppression, backlight compensation, automatic iris and many other features, there is this 2.4 megapixel high-definition image quality to meet the demand for real-time monitoring of the remote of three-dimensional warehouses. After collecting the warehouse monitoring pictures, the data needs to be feature extracted, and the deep confidence network DBN, as a typical deep learning model, has a very powerful feature extraction capability [14-15]. DBN is able to automatically learn and extract high-level abstract features of objects through multilayer nonlinear mapping, thus overcoming the limitations of traditional methods. The deep confidence network DBN contains one input layer, one output layer, and multiple hidden layers, where the DBN model consists of multiple restricted Boltzmann machine networks superimposed together, and RBM is a random generative neural network that can learn probability distributions through the input data set.By constructing a deep confidence network model and using a large amount of stereo depot image data for training, we will extract high-quality and By building a deep confidence network model and training with a large amount of stereo warehouse image data, we will extract high quality and differentiated stereo warehouse features. Fig. 2 shows the quality extraction model of stereo warehouse features by deep confidence network. In order to abstract the simple features, the weights of each layer of DBN are adjusted and updated by an unsupervised pre-training model. Through a layer-by-layer training and layer-by-layer greedy pre-training approach, the deep confidence network is able to learn high-level feature representations of image data. A specific deep confidence network structure is used and trained for stereo depot image data.

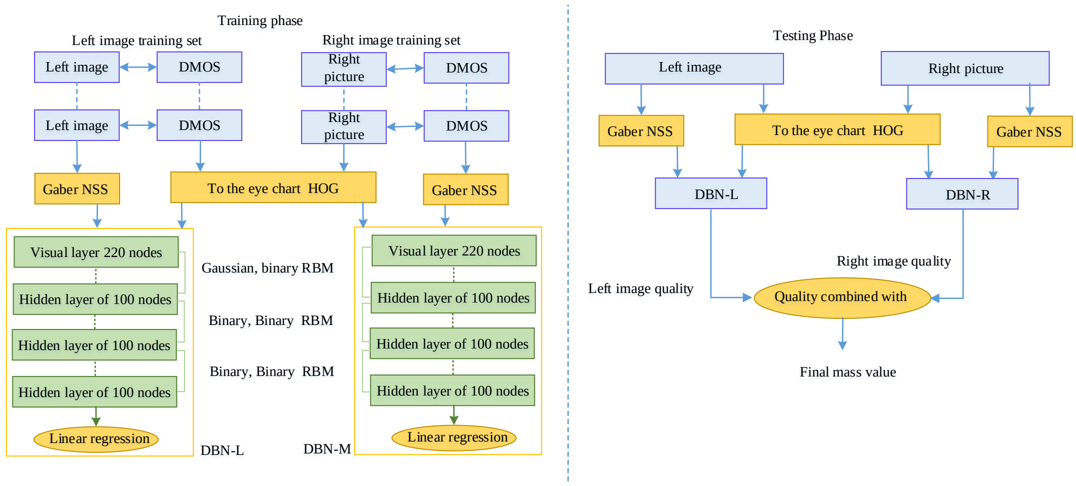


Fig. 2 Deep learning-based feature extraction model for three-dimensional warehouse

The input to the visual layer is the left image feature or the right image feature X. The final output layer is denoted by h^1,h^2,h^3 as the regression feature extraction layer, respectively, and the joint probability distribution expression of the visual and hidden layers:

$$P\left(X,h^1,h^2,h^3\right)=P\left(X\left|h^1\right.\right)\cdot P\left(h^1\left|h^2\right.\right)\cdot P\left(h^2,h^3\right) \tag{4}$$

By virtue of the learning algorithm for each neural network layer weights are trained progressively from the bottom down, with each node in the restricted Boltzmann machine network being in independent form [16]. The network probability distribution and joint distribution are formulated as follows:

$$P\left(v_i=x\left|h\right.\right)=N\left(c_i+\sigma\sum_j \omega_{i,j}h_j\right) \tag{5}$$

$$P\left(h_j=1\left|v\right.\right)=\log istic\left(b_j+\sum_i \omega_{i,j}\frac{v_i}{\sigma}\right) \tag{6}$$

$$P\left(h^1,X\right)\infty\exp\left(X^T\omega h^1+c^T h^1+b^T X\right) \tag{7}$$

$N(\cdot)$ represents the Gaussian density function, $\log istic(\cdot)$ represents the logistic equation, ω represents the network training weight, v represents the visual layer, h^i represents the hidden layer, c represents the visual bias coefficient, and b represents the hidden bias coefficient. The network architecture for training stereo storehouse image data based on deep confidence network DBN to extract high quality and discriminative stereo storehouse features is as follows:

Preprocessing Layer: this layer is mainly responsible for preprocessing the raw image data to provide the appropriate data format for the subsequent deep learning model. The preprocessing may include image cropping, scaling, normalization, etc.

DBN model body: this is the core part of the network and consists of multiple deep belief network layers. Each DBN layer usually contains multiple restricted Boltzmann machine layers. This architecture allows the network to learn more complex feature representations.

Fully Connected Layer: at the top of the DBN model, one or more fully connected layers can be added to help the model better understand and organize the learned features.

Output layer: the last layer of the network is usually the output layer, where different types of outputs can be selected depending on the needs of the specific task. For example, if the task is item categorization, then the output layer may contain one or more softmax functions for mapping the learned features to specific categories.

The main advantage of this network architecture is that it is able to learn higher-level feature representations from the original image through multiple layers of nonlinear transformations. In addition, the hierarchical structure of DBN makes it possible to capture and store useful information about the data in each layer, solve the model parameters through the input values of the regression layer, use unsupervised training to determine the weight coefficients and bias coefficients of each network layer, and combine a large amount of historical warehouse data with the DMOS units for learning and training, to automatically obtain the high-level feature representations from the original data, and then achieve the following accurate extraction of complex features within the warehouse.

3.2 Depot target detection based on hybrid Gaussian model

The development of target detection technology has made it possible to track objects in the warehouse in real time. Target detection algorithms can detect objects in the warehouse and determine their location and movement trajectory. This is very important for warehouse managers to help them understand the status and location of goods in the warehouse in real time and make timely adjustments and decisions. Hybrid Gaussian algorithm is a learning algorithm widely used in image processing and target tracking, and in this paper, we use hybrid Gaussian to model and separate the background and target of the depot, and to locate and identify the target accurately [17].

The set of historical observations of the stereo warehouse appearance image sequence is established as $X = \{X_1, X_2, \cdots, X_i\}$, the parameters of Gaussian distribution are set as $\theta_k = \{\mu_k, \sigma_k\}, k = 1, 2, \cdots, K$, the set of total parameters of the stereo warehouse appearance image, and the characterization of the probability density function of the sampled points of the stereo warehouse image as:

$$\eta\left(X\left|k,\theta_k\right.\right)=\frac{1}{\left(2\pi\right)^{n/2}\left|\sum k\right|^{1/2}}e^{-\frac{1}{2}\left(X-\mu_K\right)^T\sum^{-1}k\left(X-\mu_k\right)} \tag{8}$$

t represents the current moment, K represents the number of Gaussian distribution, x_t represents the stereo warehouse acquisition image, when the higher the number of Gaussian score, the more can indicate the stereo warehouse scene environment more chaotic and complex, combined with the actual warehouse situation, Gaussian distribution K value is taken as 5, $\omega_k\left(\sum_{i=1}^{k}\omega_i=1\right)$ represents the Gaussian distribution weight coefficient, μ_k、$\sum k$ represents the Gaussian distribution mean square deviation matrix.

The maximum expectation algorithm EM can solve non-complete data and can perform accurate parameter estimation for stereo warehouses with complex environments [18]. Using iterative maximization of data for computation, the expectation of the logarithmic function of the complete data is calculated and then the log-likelihood function is determined using the maximized expectation.

Gaussian distribution values using high-low relational equations for Gaussian distribution background generation:

$$\rho_{i,j}=\frac{\omega_i}{\sigma_i} \tag{9}$$

Where, T_1 represents the threshold value and takes the value in the range of $\left(0<T_1<1\right)$. When the value of T_1 is smaller, the value of B is higher, indicating that the Gaussian distribution is wider and the background Gaussian pixels of the stereo depot are obtained.

In order to make the algorithm better adapt to the changes in the environment of the stereo warehouse, the parameters are updated, and the weight coefficients of the Gaussian distribution are updated with the expressions:

$$\omega_{i,t}=\left(1-\alpha\right)\omega_{i,t-1}+\alpha M_{i,t} \tag{10}$$

In the formula, α represents the deep learning rate of the weights, and the size of α determines the Gaussian background update efficiency, if the background pixel can be matched with any Gaussian distribution value, i.e. $M_{i,t}=1$. If the background pixel cannot be matched with any Gaussian distribution value, i.e. $M_{i,t}=0$. Gaussian mean distribution:

$$\mu_{i,t} = (1-\beta)\mu_{i,t-1} + \beta X_{i,t} \tag{11}$$

Gaussian variance distribution:

$$\sigma_{i,t}^2 = (1-\beta)\sigma_{i,t-1}^2 + \beta\left(x_{i,t} - \mu_{i,t}\right)^T\left(x_{i,t} - \mu_{i,t}\right) \tag{12}$$

Expressions for the learning rate of Gaussian mean distribution and Gaussian variance distribution:

$$\beta = \alpha\eta\left(x_{i,t} \middle| \mu_{i,t}, \sigma_{i,t}\right) \tag{13}$$

In the formula, the learning rate β of the mean and variance is changed by the change of the weight learning rate α .

Matching calculation for the mean value of the Gaussian distribution of the pixel values $x_{i,t}$ of the image of the t th stereo depot:

$$\left|x_{i,t} - \mu_{j,t}\right| < T\sigma_{j,t} \tag{14}$$

In the formula, $\mu_{j,t}$ stands for Gaussian mean and $\sigma_{j,t}$ Gaussian variance, and the matching effect is optimal when the threshold T_1 is 2.5. When the pixel value of the stereo depot image satisfies the formula (14), it means that the matching is completed and the pixel point at this moment is the background point of the center of the depot. If the Gaussian variance does not satisfy Equation (14), a new Gaussian distribution value needs to be introduced to replace the weight minimum coefficient using the new Gaussian distribution. Other Gaussian models need to keep the same size of mean and variance, and the Gaussian weight decay formula:

$$\omega_{i,t} = (1-\alpha)\omega_{i,t-1} \tag{15}$$

The weight values are normalized to update the iterative stereo warehouse background image with the following weight normalization expression [19]:

$$\omega_{i,j,xy}^{*} = \frac{\omega_{i,y,xy}}{\sum_{j=1}^{K}\omega_{j,t,xy}} \cdots \tag{16}$$

At this moment, the current pixel point is the foreground pixel point of the stereo warehouse. Fig. 3 shows the flow chart of the stereo warehouse target detection based on the hybrid Gaussian model.

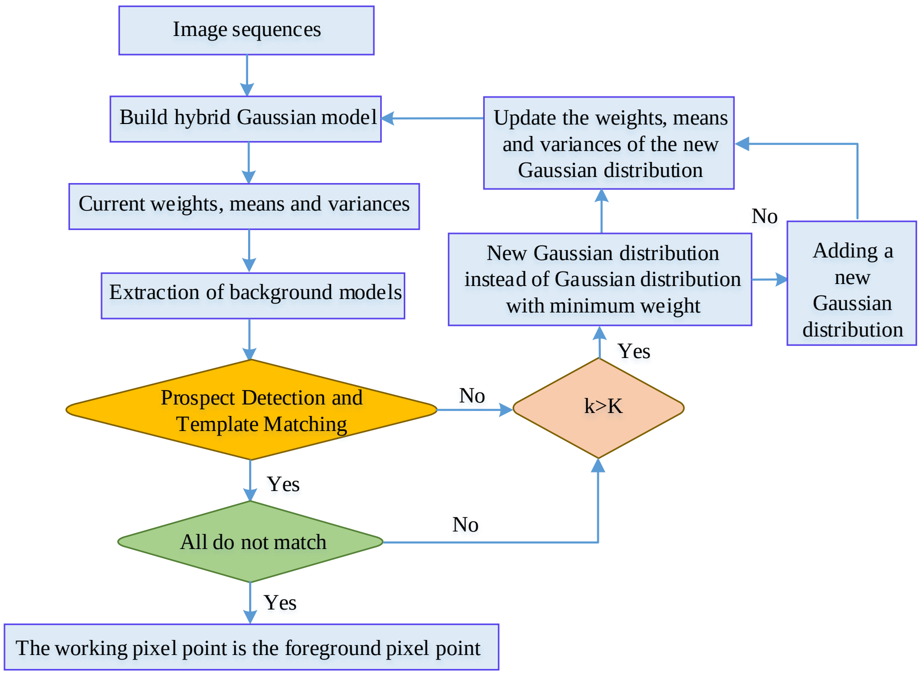


Fig. 3 Flowchart of stereo warehouse target detection based on hybrid Gaussian model

3.3 Kalman filter-based real-time monitoring and tracking of three-dimensional warehouses

Kalman filtering is an algorithm that uses a linear system state function to optimally estimate the state of a target object by virtue of the input and output observations [19-20]. The Kalman filter-based approach can predict and correct the state of the object using a priori and measurement information, thus improving the accuracy and stability of object tracking. Therefore, the Kalman filter algorithm is utilized for real-time monitoring and tracking of the stereo warehouse.

Target monitoring is mainly through the target feature set, predicting the target matching range in the next time period, and using the results of the stereo depot target matching features as the input values of the Kalman filter in the next time period. Establishing linear measurement equations, the state variables of the Kalman system characterize the center of mass and velocity of the stereo depot, and the expressions of the state representation and observed representation of the stereo depot:

$$X(k)=\left[x(k),y(k),W(k),H(k),x'(k),y'(k)\right] \tag{17}$$

$$Z(k)=\left[x(k),y(k),w(k),h(k)\right]^{T} \tag{18}$$

In the equation, $X(k)$ represents the cubicle state representation, $Z(k)$ represents the cubicle observation representation, $x(k)$ represents the horizontal coordinate of the center of mass, $y(k)$ represents the vertical coordinate of the center of mass target, $W(k)$ represents the width of the cubicle rectangle, $H(k)$ represents the height of the cubicle rectangle, $x'(k)$ represents the horizontal coordinate of the center of mass velocity, and $y'(k)$ represents the vertical coordinate of the center of mass velocity.

Due to the small movement time Δk between two adjacent images, a uniform movement is needed in time interval Δk. The expression of the depot state matrix A and the observation matrix H:

$$A(k,k-1)=\begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}\cdots \tag{19}$$

$$H(k)=H=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}\cdots \tag{20}$$

The system updates the covariance matrix [20]:

$$P(0|0)=\begin{bmatrix} 15 & 0 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 & 0 \\ 0 & 0 & 15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 15 & 0 \\ 0 & 0 & 0 & 0 & 15 \end{bmatrix}\cdots \tag{21}$$

The value of the tracking window (x, y) for the motion monitoring time Δk can be calculated with the equation :

$$\begin{aligned}
&S_x(k+1)=1.5*w(k+1) \quad S_y(k+1)=1.5*h(k+1)\\
&x(k+1)=x(k)+\Delta x \qquad\qquad y(k+1)=y(k)+\Delta y\cdots\\
&x(k+1)-S_x(k+1)/2\le x\le x(k+1)-S_x(k+1)/2\cdots\\
&y(k+1)-S_y(k+1)/2\le y\le y(k+1)-S_y(k+1)/2\cdots
\end{aligned} \tag{22}$$

Extracting target features and Kalman monitoring and tracking respectively, the steps of remote real-time monitoring and tracking of the three-dimensional warehouse are shown in Fig. 4.

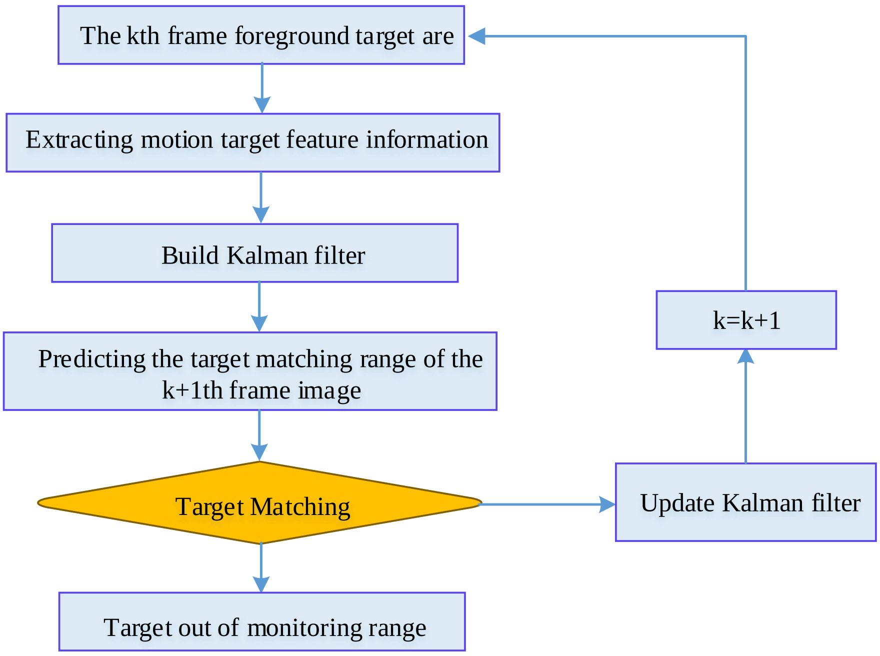


Fig. 4 Remote real-time monitoring and tracking process of the cubicle

The steps to monitor the tracking of the depot using Kalman filter are as follows:

(1) Using the center of mass of the target, the width and height of the external rectangle and the velocity of the center of mass, the characteristics of the moving target of the three-dimensional depot are determined.

(2) In the initial stage of tracking, the center of mass and velocity of the target are unknown, and the feature vector initialization is set to 0.

(3) Monitor the target area of the stereo depot. The range of the target in the next frame is predicted by Kalman filter, and the real-time monitoring of the moving target is performed within the range of the predicted depot.

4. Experimental results and analysis

4.1 Experimental background

In order to verify the performance of the proposed method for remote real-time monitoring and tracking of three-dimensional depots, a regional power grid material depot at all levels at the provincial level is selected as the experimental sample. The warehouse of the experimental regional power grid company has many miscellaneous objects and inefficient logistics materials management. Based on the MATLAB2016 programming experimental platform, the monitoring performance and tracking performance are verified using Matlab programming on a PC with 64GB memory. In terms of data pre-processing, the acquired image data were pre-processed, including image denoising, image enhancement and image size unification steps, to improve the accuracy and effectiveness of subsequent processing. In addition, the data were annotated and calibrated with the bounding box of the depot and the location information of the objects for the algorithm training and evaluation.

4.2 Real-time remote monitoring test of three-dimensional storage room

AUC is a performance metric that measures the strength of a learner, AUC is the likelihood that given a random positive sample and a negative sample, the classifier is more likely to output that probability value of that positive sample as positive than the classifier is more likely to output that probability value of that negative sample as positive. So AUC responds to the ability of the classifier to rank the samples. Therefore, the larger the AUC value, the more effective the classifier is. In binary classification tasks, the AUC value is usually used as an evaluation criterion for the model. The library monitoring video is categorized into 10 monitoring attributes, and the average AUC results are compared with this paper's method by three other methods in response to these complex monitoring attribute tracking factors, and Table 1 shows the comparison of the average AUC results of the four methods for the complex monitoring attribute tracking factors. AUC results of the four methods for complex monitoring attribute tracking factors. The AUC results of the proposed method are optimal in the complex monitoring attributes of camera motion, fast motion, similar target interference, occlusion interference, target out-of-field, illumination change, background change, viewpoint change, low resolution, and scale change, and the AUC results are as high as 0.93 in the target out-of-field. The optimal AUC results of the greedy algorithm method are 0.76 in the occlusion interference of the complex monitoring attributes. The optimal AUC result of artificial neural network in similar target interference of complex monitoring attributes of the offset tracking method is 0.68, which is average and not ideal, while the monitoring effects of complex monitoring attributes of the proposed method are above 0.85, which indicates that the proposed method can obtain the information of items and environment inside the warehouse based on overcoming problems such as light changes and target occlusion, and can achieve through the network monitoring and management of remote storage.

In order to further illustrate the applicability of the proposed method for monitoring, the item status of the three-dimensional storage room is monitored remotely in real time, and the monitoring accuracy and monitoring time consumption are compared and analyzed for the change in the number of items in the storage room, the change in the location of the items in the storage room, the change in the outgoing items in the storage room, and the change in the incoming items in the storage room. Fig. 5 shows the comparison results of monitoring the change of the state of the items in the three-dimensional warehouse of the four methods.

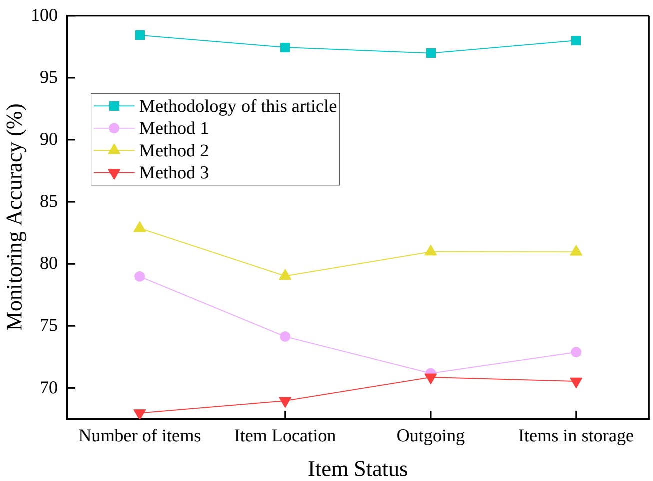
Table 1

AUC results for complex monitoring of attribute tracking factors

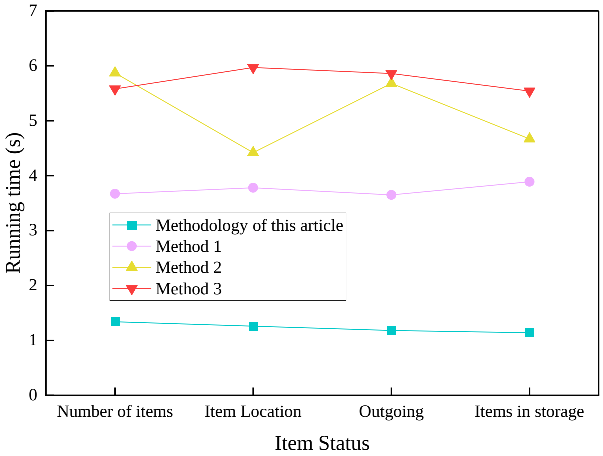
Monitoring video attributes	Total number of videos	Proposed method	Greedy algorithm	Offset tracking method	Artificial neural network
Camera motion	12	0.86	0.65	0.75	0.59
Fast motion	10	0.87	0.67	0.73	0.56
Similar target interference	8	0.88	0.71	0.69	0.68
Occlusion interference	13	0.91	0.76	0.68	0.66
Target out of view	17	0.93	0.67	0.70	0.53
Illumination change	10	0.86	0.68	0.66	0.52
Background change	11	0.85	0.64	0.71	0.59
View angle change	8	0.90	0.64	0.69	0.57
Low resolution	12	0.85	0.72	0.68	0.61
Scale change	15	0.85	0.65	0.69	0.58

Fig. 5(a) shows the accuracy of monitoring the changes in the state of warehouse items. The proposed method has an accuracy of 98.43 in monitoring the changes in the quantity of warehouse items, the greedy algorithm has an accuracy of 97.45 in monitoring the changes in the location of warehouse items, the offset tracking method has an accuracy of 96.98 in monitoring the outgoing of warehouse items, and the artificial neural network has an accuracy of 98.01 in monitoring the incoming of warehouse items. The optimal monitoring accuracy of the greedy algorithm is only 78.98, the optimal monitoring accuracy of the offset tracking method is only 82.87, and the optimal monitoring accuracy of the artificial neural network is only 70.86. This shows that the proposed method can monitor and track the location and status of the items in the warehouse in real time, and realize the dynamic management of the items in the warehouse. The proposed method can improve the overall management efficiency and operation level of the warehouse. This is due to the fact that the proposed method uses the hybrid Gaussian model to model and separate the complex background and target of the three-dimensional warehouse, and successfully detects the target of the three-dimensional warehouse accurately, which can not only detect the objects in the warehouse and calculate the position and motion trajectory of the three-dimensional warehouse.

Fig. 5(b) shows the running time for monitoring the changes in the state of the items in the warehouse. The minimum monitoring time of item state change for the greedy algorithm is 3.65s, the minimum monitoring time of item state change for the offset tracking method is 4.42s, and the minimum monitoring time of item state change for the artificial neural network is 5.54s, indicating that the proposed method can complete the monitoring of item state change in the shortest running time.



(a) Accuracy of monitoring changes in the status of warehouse items



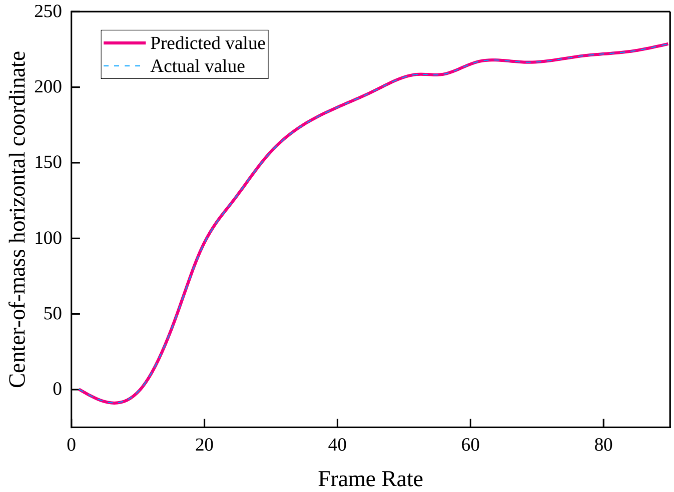
(b) Running time for monitoring changes in the status of warehouse items

Fig. 5 Comparison of the monitoring effect of object state change

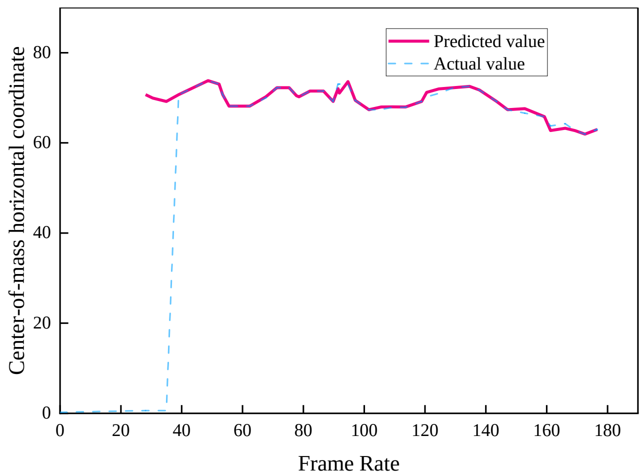
4.3 Experimental results of remote tracking in the three-dimensional warehouse

To further validate the practical application potential and advantages of the proposed technology and to verify the intelligence and efficiency of the warehouse management. The results of the three-dimensional depot remote tracking experiment are shown in Fig. 6. Fig. 6(a) shows the predicted value of the depot mass center transverse coordinate compared with the real value. The predicted value of the depot mass center coordinate fits the curve of the real value of the depot mass center coordinate very well, when the number of frames is 20, the predicted value of the depot mass center transverse coordinate is at 100, which is equal to the predicted value of the actual transverse coordinate. Still near the end, at 80 frames, the predicted value of the depot mass coordinate is at 210, which is also equal to the predicted value of the actual horizontal coordinate.

Fig. 6(b) shows the comparison results between the predicted and real values of the depot mass center longitudinal coordinates, and the predicted values are able to approximate the real values very well. At the beginning, when the frame number is 13, the predicted value of the depot mass vertical coordinate is at 141 and the predicted value of the actual vertical coordinate is 149. When the frame number gradually increases, the fit between the two is better and better, and when the frame number exceeds 50, the predicted value of the depot mass horizontal coordinate and the predicted value of the actual horizontal coordinate have no error.



(a) Predicted versus true values of the depot mass center transverse coordinates



(b) Comparison results of the longitudinal coordinates of the depot mass center

Fig. 6 Comparison of predicted and true values of coordinates of the cubicle

The difference between the predicted and true values of the center of mass coordinates of the cubic depot is shown in Fig. 7, and it can be analyzed that the circles indicate the difference between the horizontal coordinates of the predicted values and the horizontal coordinates of the true values, and the triangles indicate the difference between the vertical coordinates of the predicted values and the vertical coordinates of the true values. As the number of frames increases, the error between the predicted value and the true value becomes smaller and smaller, which indicates that the prediction ability of the cartesian method gradually stabilizes with the operation time. The operation speed of the whole system can reach more than 15fps, which can achieve the real-time requirement for accurate tracking of the three-dimensional warehouse.

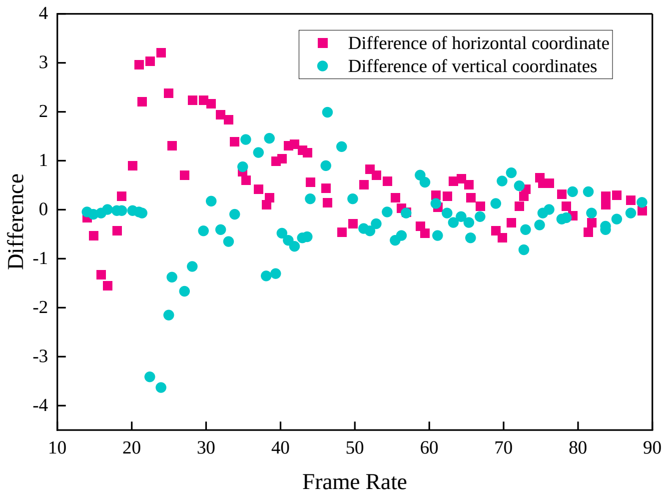


Fig. 7 Difference between predicted and true center-of-mass coordinates

5. Discussion

Compared to previous research works, the study in this paper reveals potential advances in deep learning and target detection based algorithms, which can significantly improve the accuracy and efficiency of detection and tracking. First, the proposed study is consistent with the conclusions drawn from the literature [11], which emphasizes the importance of designing more efficient neural network structures. Based on the literature [11], future research on novel architectures is suggested to strike a balance between computational efficiency and detection performance. Secondly, the study confirms the findings of literature [12], emphasizing the need for target detection algorithms. Based on these contributions, future research could explore literature [14] and [15] wireless network domain technologies to provide intelligent decision support for inventory management, logistics scheduling, and anomaly detection by analyzing and mining warehouse monitoring and tracking data in real time.

6. Conclusion

Traditional warehouse monitoring and tracking technology cannot meet the needs of modern warehouse management, so the remote real-time monitoring and tracking technology of three-dimensional warehouse based on deep learning and target detection has become a research technology that has attracted much attention. Through the wireless sensor nodes to the three-dimensional warehouse all-round coverage, by virtue of the depth of the confidence network on the three-dimensional warehouse image data for automatic learning, to determine the complex characteristics of the warehouse interior and high-level three-dimensional warehouse features. Using hybrid Gaussian model to identify the target location of the warehouse, and finally combining Kalman filtering to introduce a priori information and measurement information to predict and correct the state of the object to complete the real-time monitoring and tracking of the three-dimensional warehouse. Experimental results: the complex monitoring attribute AUC value of the proposed method is as high as 0.93, and the monitoring accuracy is as high as 98.43%, which is able to detect the objects in the warehouse and calculate the position and motion trajectory of the three-dimensional warehouse. And the monitoring time consumed in the warehouse items into the warehouse is only 1.14s, which greatly improves the real-time monitoring and efficiency. At the same time, the predicted values of the center of mass coordinates of the warehouse fit the real values of the actual coordinates excellently, and the running speed of the whole system can reach more than 15 fps, which verifies the tracking ability of the proposed method for the goods in the warehouse.

However, there are still some challenges and problems, the method in this paper is limited by the accuracy and robustness of feature extraction, the problem of easy misdetection and omission for the warehouse environment with complex backgrounds and lighting variations, and the processing and analysis of large-scale data, which need to be further researched and solved. Therefore, this thesis intends to conduct an in-depth study on the remote real-time monitoring and tracking technology for three-dimensional warehouses in order to solve the problems in practical applications. Future research can address these issues in depth to further improve the performance and practicality of the three-dimensional warehouse monitoring and tracking technology. In addition, other advanced deep learning models can be explored to be applied to three-dimensional warehouse monitoring to improve the accuracy and robustness of detection and tracking.

R E F E R E N C E S

[1] Nee, A., & Chamkha, A. J. (2023). A simplified gpu implementation of the hybrid lattice boltzmann model for three-dimensional high rayleigh number flows. International Journal of Applied Mechanics, 15(06).

[2] Xu, B., & Ren, Q. (2023). Interfaces with boundary intersection for an inhomogeneous allen-cahn equation in three-dimensional case. Communications on Pure and Applied Analysis, 22(5), 1585-1628.

[3] En Vandana. (2023). Analysis of two-warehouse effects on inventory models under preservation technology and ramp-type demand with complete and partial backlogging. Journal of Industrial and Management Optimization, 19(11), 8089-8113.

[4] Hao, Q., Wang, Z., & Qin, L. (2019). Design of beidou satellite system in ocean logistics real-time tracking system. Journal of Coastal Research, 94(sp1), 204.

[5] Li, Y., Shi, X., Diao, H., Zhang, M., & Wu, Y. (2021). Optimization of warehouse management based on artificial intelligence technology. Journal of Intelligent and Fuzzy Systems, (6), 1-8.
[6] Liu, B., Huang, L., Zhang, R., & Hu, C. (2023). Strategic outsourcing decisions of new entrant and competing incumbent manufacturer in a supply chain with common supplier. Journal of Industrial and Management Optimization, 19(8), 5842-5868.
[7] He, F., Xu, J., Zhong, J., Chen, G., & Peng, S. (2021). Optimal sensor association and data collection in power materials warehouse based on internet of things. Energies, 14.
[8] Xue, T., Wang, S., Han, F., & Sun, D. (2020). Anti tracking technology of urban water supply pollution source based on water quality monitoring. Journal of Coastal Research, 104(sp1).
[9] Zhao, G., Wang, L., Deng, K., Wang, M., & Luo, Q. (2021). An adaptive offset-tracking method based on deformation gradients and image noises for mining deformation monitoring. Remote Sensing, 13(15), 2958.
[10] Hu, Y., Xiao, M., Zhang, K., Kong, Q., & Ge, P. (2020). Aerial infrared target tracking in severe jamming using skeletal tracking technology. Infrared Physics & Technology, 103545.
[11] Sun, X., Huang, M., Liu, Y., & Zha, X. (2018). Investigation of artificial neural network algorithm based igbt online condition monitoring. Microelectronics Reliability.
[12] Abdullah, A. N., & Ali, L. (2021). Comparative study of super-performance doa algorithms based for rf source direction finding and tracking. Technology Reports of Kansai University, 63(2), 7283-7294.
[13] Zhou, N., Lau, L., Bai, R., & Moore, T. (2021). A genetic optimization resampling based particle filtering algorithm for indoor target tracking. Remote Sensing, 13(1), 132.
[14] Kagi, S., & Mathapati, B. S. (2022). Optimal trained deep neural network for localization in wireless sensor network. International Journal of Computational Methods.
[15] Zhiyao, L., Siru, O., Xiaofeng, G., & Guihai, C. (2021). Two-hop relay deployment based on user trajectory in wireless networks. The Computer Journal, (12), 12.
[16] Chen, Y. H., Lai, Y. C., Jan, P. T., & Tsai, T. Y. (2021). A spatiotemporal-oriented deep ensemble learning model to defend link flooding attacks in iot network. Sensors, 21(4), 1027.
[17] Hua, J., & Tao, S. (2022). Evaluation algorithm of skilled talents' quality based on deep belief network model. Journal of Interconnection Networks.
[18] Fu, Q. L., Fujii, M., & Ma, R. (2023). Development of a gaussian-based alignment algorithm for the ultrahigh-resolution mass spectra of dissolved organic matter. Analytical chemistry, 95(5), 2796-2803.
[19] Envelope, W. L., & Envelope, J. Z. P. (2022). Clt for spiked eigenvalues of a sample covariance matrix from high-dimensional gaussian mean mixtures. Journal of Multivariate Analysis.
[20] Li, H., Geng, J., & Xie, J. (2022). Robust adaptive beamforming based on covariance matrix reconstruction with rcb principle. Digital Signal Processing.