


Research Article

Radar Target Recognition and Location Based on CapsNetv2

Jiaying Hao ¹, Xuettian Wang,¹ Sen Yang,² and Hongmin Gao¹

¹School of Information and Electronics, Beijing Institute of Technology, No. 5 zhongguancun South Street, Beijing 100081, China

²Department of UAV Engineering, Army Engineering University, 97 Heping West Road, Shijiazhuang 050003, China

Correspondence should be addressed to Jiaying Hao; 873328461@qq.com

Received 29 March 2022; Accepted 14 May 2022; Published 26 May 2022

Academic Editor: Maode Ma

Copyright © 2022 Jiaying Hao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For precise detection and positioning of weapons and equipment under complex ground backgrounds and weather-changing aerial backgrounds. Compared with the traditional convolutional neural networks, the Capsule Network (CapsNet) is more suitable for identifying weapons and equipment in complex backgrounds because it uses vectors as input for the first time, which can well retain the characteristic information such as the direction and the angle of the target. Therefore, this paper proposes a radar target classification algorithm based on the combination of CapsNetv2 and infrared lidar, which simplifies the convolutional layer of the traditional 9×9 capsule network through a 1×1 reduction layer and a 3×3 convolution kernel, and adopts a double-layer capsule layer. Two prediction frames are obtained to improve the recognition accuracy; at the same time, the output volume retains the direction and the angle, which can more accurately classify the radar targets in various complex backgrounds. Applying the method proposed in this article to the MSTAR dataset shows that the radar target positioning is accurate. The rate increases to 99.5%. Finally, compared with the AlexNet and the YOLOv4 methods designed by Alex Krizhevsky, the proposed radar target recognition method can accurately and quickly identify weapons and equipment from complex backgrounds. The results obtained from the CapsNetv2 are accurately compared with other methods' in complex backgrounds. The proposed method significantly improves the efficiency of military inspections.

1. Introduction

Radar technology plays an important role in the field of modern target detection. It is widely used in military and civil transportation fields due to its all-weather and omnidirectional work characteristics. Target recognition is one of the basic tasks in computer vision. Identifying the target area and obtaining the accurate position of the target lay the foundation for the next information processing of the carrier and improve the perception ability of machine recognition. The current main model is to process the visible light image accordingly. However, the visible light image is susceptible to environmental lighting. Under low light, dark or shadow conditions blocked by surrounding interference, the processing data becomes more complicated. To achieve high-reliability classification and recognition effects, modern pattern recognition theories and methods are usually used for classifier design, such as statistical-based pattern classification methods, feature extraction methods, and neural network-based pattern classification methods.

Statistics-based classification and recognition algorithms use probability models to obtain the feature vector distribution of each category and classify the unknown samples. For example, Shen Yanyan obtained the likelihood function by extracting the ocean wave radar echoes and used Bayesian classifiers for classification [1]. Liu Jingrui and others established a weather radar warning system using probability statistical models to distinguish between strong and weak rainfalls [2]. However, the weather environment is complex and changeable, and there are many interference factors, making the actual task of processing the radar signals much more difficult than processing and identifying the visible light images.

The traditional feature methods mainly match the known features by extracting the feature points of the target. Commonly used feature matching methods include the histogram of oriented gradient (HOG) feature [3], the scale-invariant feature transform (SIFT) feature [4], and the speed-up robust features (SURF) feature [5]. In 2001, the American company ENSCO developed the Visual Identity

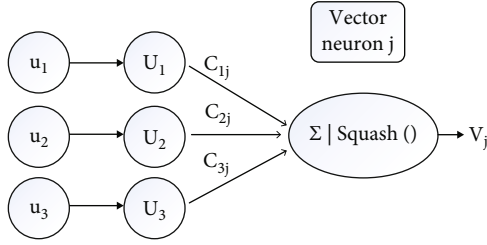


FIGURE 1: Capsule neuron model.

System (VIS) track video detection system to realize real-time detection of the working status of PandaPal fasteners [6]. In 2005, the German railway engineering company GBM Wiebe developed the GeoRail-Xpress comprehensive inspection vehicle that was able to perform a real-time inspection of the entire railway electrical equipment system [7]. However, because it is necessary to extract and classify multiple regions of the image, the recognition speed was slow, and it was difficult to meet the requirements of real-time detection.

In recent year, the target recognition algorithms based on deep learning have made significant progress compared with the traditional target detection algorithms. The representative algorithms include R-CNN [8], Fast-RCNN [9] and Faster-RNN [10]. However, the detection steps of these methods are more complicated, and the real-time effect is poor. The AlexNet [11] and YOLO [12] that have appeared one after another can meet the requirements of real-time detection, but often require a lot of data for neural network training, and the training weights are easy to overfit. Moreover, the technical requirements for the equipment are relatively high. In 2017, Hinton proposed that the Capsule Network, referred to as CapsNet, would possibly replace the traditional CNN network, bringing new opportunities to the field of deep learning [13]. For example, the literature [14] used the capsule network to classify handwritten digits. Due to the single characteristics of the digits, the recognition rate is high. However, the radar targets are generally weapons and equipment with complex structures and are easily affected by the conditions such as illumination and angle.

This paper proposes a radar target recognition and location algorithm based on CapsNetv2. For the characteristics of weapons and equipment, light intensity, position, deformation, angle, texture, and position information should be considered. Therefore, these six features are selected as the input vectors, and then a 1×1 reduction layer combined with a 3×3 convolutional layer is used to simplify the traditional Capsule 9×9 capsule neurons. Then the MSTAR dataset is trained and learned through the double-layer capsule network, and two prediction boxes are obtained, one large and one small, to complete the recognition under different complex backgrounds. Finally, the improved CV model and the infrared Lidar uses edge detection to accurately locate the location of hazardous enemy weapons and equipment.

2 Structure. A paper for publication can be subdivided into multiple sections: title, list of all the authors and their affiliations, a

concise abstract, keywords, main text (including figures, equations, and tables), acknowledgement, references, and appendix.

2. Capsnetv2

In 2011, Hinton proposed the concept of capsule [15]. Unlike the traditional scalar neurons, the capsule network is a vector composed of many neurons. The vector length of the capsule neuron model indicates the possibility of the existence of the target passed by the upper network, and its direction represents the actual state of the entity, that is, "Instance parameters", as shown in Figure 1 [16].

The dynamic routing algorithm (Squash) solves the problem by the output value of the capsule.

The update formula is:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (1)$$

$$b_{ij} = b_{ij} + U_i \cdot V_j \quad (2)$$

where c_{ij} is the dynamic routing coupling coefficient and k is the number of initial similarity weights b_{ij} .

The capsule output S_j is obtained from the lower-level capsule inputs U_i and c_{ij} :

$$S_j = \sum_i c_{ij} U_i \quad (3)$$

where U_i is derived from $U_i = W_{ij} u_i$,

W_{ij} is the weight of the capsule network.

The output V_j should be expressed as a probability. Thus, the output value should be controlled between [0, 1], which can be obtained by nonlinear compression:

$$V_j = \frac{\|S_j\|^2}{1 + \|S_j\|^2} \times \frac{S_j}{\|S_j\|} \quad (4)$$

The principle of CapsNetv2 is roughly the same as that of the capsule network. The image is first input to the convolutional layer (ReLU), and a basic capsule layer is obtained through the convolution operation. Then the data of the basic capsule layer is transmitted to the image through the dynamic routing algorithm (squash). The capsule layer then transfers the image capsule layer data to the feature capsule layer, and finally uses the fully connected layer to reorganize and model the feature capsule layer data. However, the CapsNetv2 consists of two image capsule layers and two feature capsule layers. If training is performed when the data of one layer of the capsule layer has over-fitting, it can ensure the success of the training of the other capsule layer. The structure of the CapsNetv2 is shown in Figure 2.

The radar target image is composed of 3 categories, which are set as BTR70(armored transport vehicle), BMP2(infantry fighting vehicle), and T72(tank). The moduli of the three types of target vectors are calculated and the

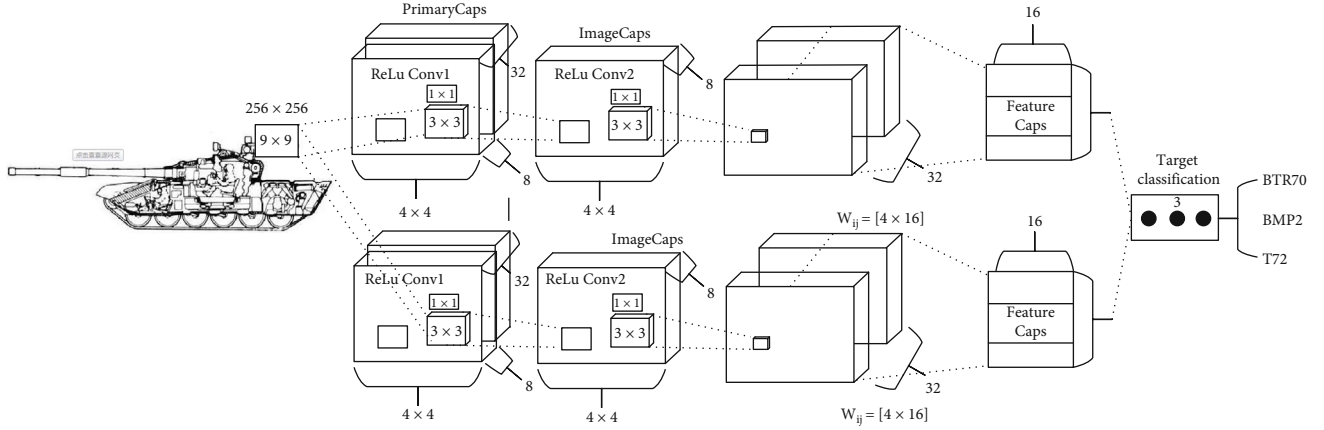


FIGURE 2: Schematic diagram of the CapsNetv2 structure.

TABLE 1: Performance comparison of different models.

Model	Top-1%	Top-5%	GPU/ms	CPU/S
AlexNet	57.0	90.0	2.5	0.30
Yolov4	63.0	92.0	2.7	0.26
CapsNet	70.0	92.5	1.8	0.66
CapsNetv2	72.0	93.5	1.3	0.18

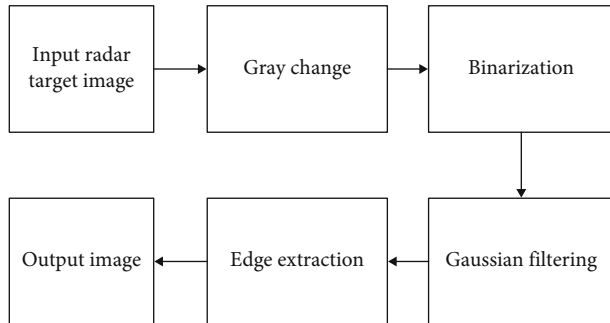


FIGURE 3: Flow chart of hanging string image pre-processing.

vector with the largest modulus value is the category with the highest possible target probability.

The AlexNet, Yolov4 and the traditional capsule network are used to compare the performance of the CapsNetv2 and classify the image dataset. Table 1 compares the Top-1% and Top-5% classification performance of each model for the same dataset, where the GPU model is Titan X, and the CPU model is Intel I7-10700(4GHz).

As shown in Table 1, CapsNetv2 has higher classification accuracy in Top-1% and Top-5% compared with AlexNet, Yolov4 and CapsNet. Moreover, the recognition time of GPU and CPU is less, indicating that the CapsNetv2 has better performance.

3. Principles of Radar Target positioningSubheadings

3.1. Radar Image Preprocessing. In the radar target recognition technology, the collected radar image contains various disadvantages such as noise, jitter, and weak light due to

its complex background, weather and other factors that will affect the model training and recognition results. Therefore, it is necessary to pre-process and correct the collected original image and then extract the feature value of the target and separate the target from the background.

The pre-processing steps include grayscale change, binarization, noise reduction, filtering and edge extraction. The specific flow chart of pre-processing is shown in Figure 3.

- (1) Perform grayscale processing on images of different categories in the mSTAR dataset. The results are shown in Figure 4(a);
- (2) Binarize the grayscale processed image to remove the influence of complex background, that is, set the pixel point to 0 or 255, where the target gray value is 255, and the other background is 0 as shown in Figure 4(b);
- (3) Noise will reduce the quality of the image, and the collected radar target image is usually accompanied by auxiliary equipment and anti-jamming equipment that contain a lot of Gaussian noise. Therefore, this article uses Gaussian filtering to process the image, as shown in Figure 4(c);
- (4) By comparing the radar target recognition effect with Robert, Sobel or LOG operator, the edge of the target detected by the Canny algorithm is more complete. Therefore, this paper uses the Canny algorithm to extract the edge of the radar target as shown in Figure 4(d).

3.2. Improved CV Model. The Chan-Vese(CV) model is used to divide the fuselage and barrel of the T72 tank. The energy function of the CV model is [17]:

$$E(C, c_1, c_2) = \mu \cdot L(C) + \lambda_1 \int_{i(C)} (\mu_0(x, y) - c_1)^2 + \lambda_2 \int_{i(C)} (\mu_0(x, y) - c_2)^2 \quad (5)$$

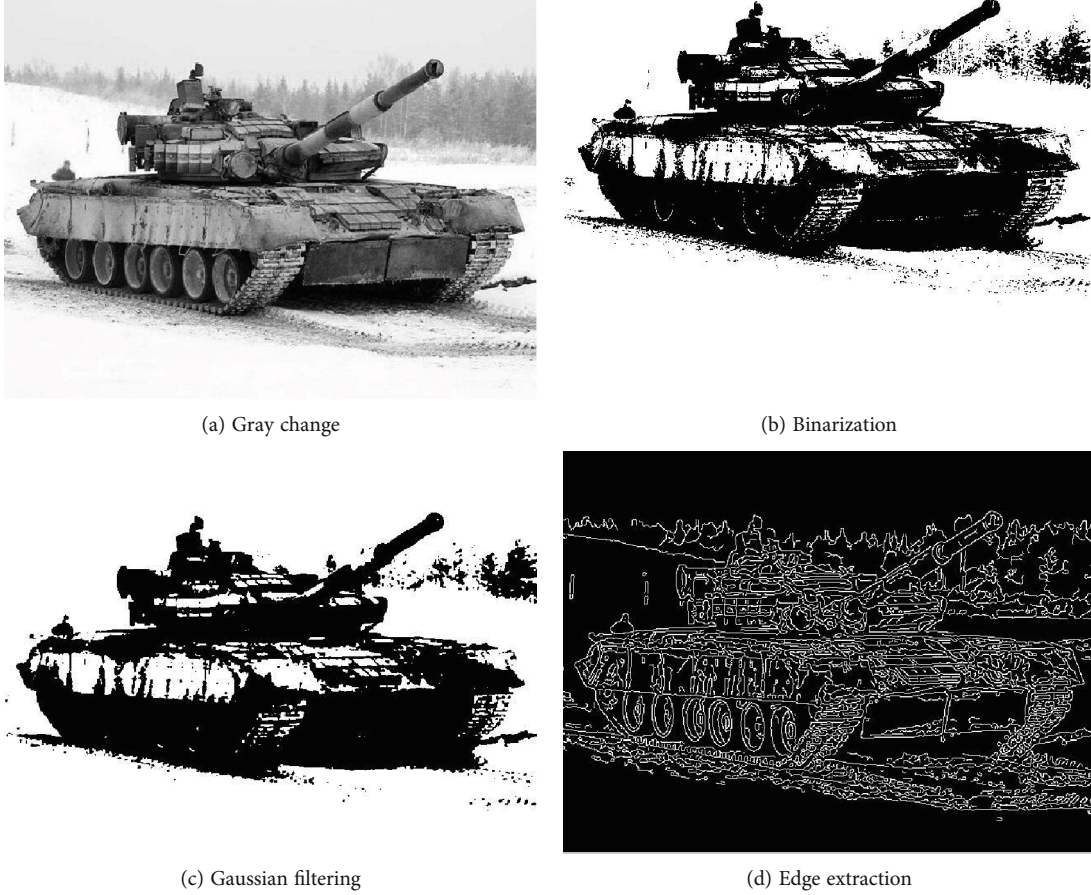


FIGURE 4: Processing results of the T72 tank.

$$c_1 = \frac{\int_{\Omega} \mu(x, y) H_{\varepsilon}(\phi) dx dy}{\int_{\Omega} H_{\varepsilon}(\phi) dx dy} \quad (6)$$

$$c_2 = \frac{\int_{\Omega} \mu(x, y) [1 - H_{\varepsilon}(\phi)] dx dy}{\int_{\Omega} [1 - H_{\varepsilon}(\phi)] dx dy} \quad (7)$$

In the formula, μ is the CV model constant; $L(C)$ is the arc length of the curve C ; $\mu \cdot L(C)$ is the length term, which can smooth the evolution curve; λ_1 and λ_2 are the weight coefficients, both greater than 0; $\mu_0(x, y)$ is the image pixel gray value; c_1 and c_2 are the average gray values of the pixels outside and inside the image evolution curve, respectively, and $H_{\varepsilon}(\phi)$ is the regularization step function.

Considering that targets such as armored vehicles or tanks are regular models with regular shapes and horizontal symmetry, adding the level set method can better correct the contour topological changes. The level set evolution Euler-Lagrangian equation is:

$$\begin{aligned} \partial\phi/\partial t = & \delta_1 \left[\mu \operatorname{div}(\nabla\phi/|\nabla\phi|) - \lambda_1 \int_{i(C)} (\mu_0(x, y) - c_1)^2 \right] \\ & + \lambda_2 \int_{i(C)} (\mu_0(x, y) - c_2)^2 \end{aligned} \quad (8)$$

where δ_1 is the global function, as the impulse function of

the CV model; div is the divergence operator and $\operatorname{div}(\nabla\phi/|\nabla\phi|)$ is the curvature of the evolution curve.

This paper selects the T72 heavy tank with obvious barrel characteristics as the segmentation object. The rectangle is set as the initial contour line through the CV model. The image needs to be corrected by Hough to accurately locate the damaged location.

Figure 5 shows the original image (a), the initial circular contour (b), the level set function (c) and the ellipse contour positioning result corrected by Hough transform (d). It can be seen that for the barrel with obvious characteristics on the tank, the rectangular initial contour of the CV model is modified by Hough change. The elliptical contour can better locate the whole part of the tank compared with the circular initial contour. Therefore, this article adopts Hough change. The revised CV model is used for the positioning of the T72 tank.

3.3. Infrared Lidar Positioning. Suppose T_{lidar}^{camera} is calibrated as the conversion matrix from the lidar coordinate system to the camera coordinate system, and the formula is as follows:

$$P_{camera}(X, Y, Z) = T_{lidar}^{camera} P_{lidar}(X, Y, Z) \quad (9)$$

According to formula (9), the relative three-dimensional coordinates of the target in the camera can be obtained.

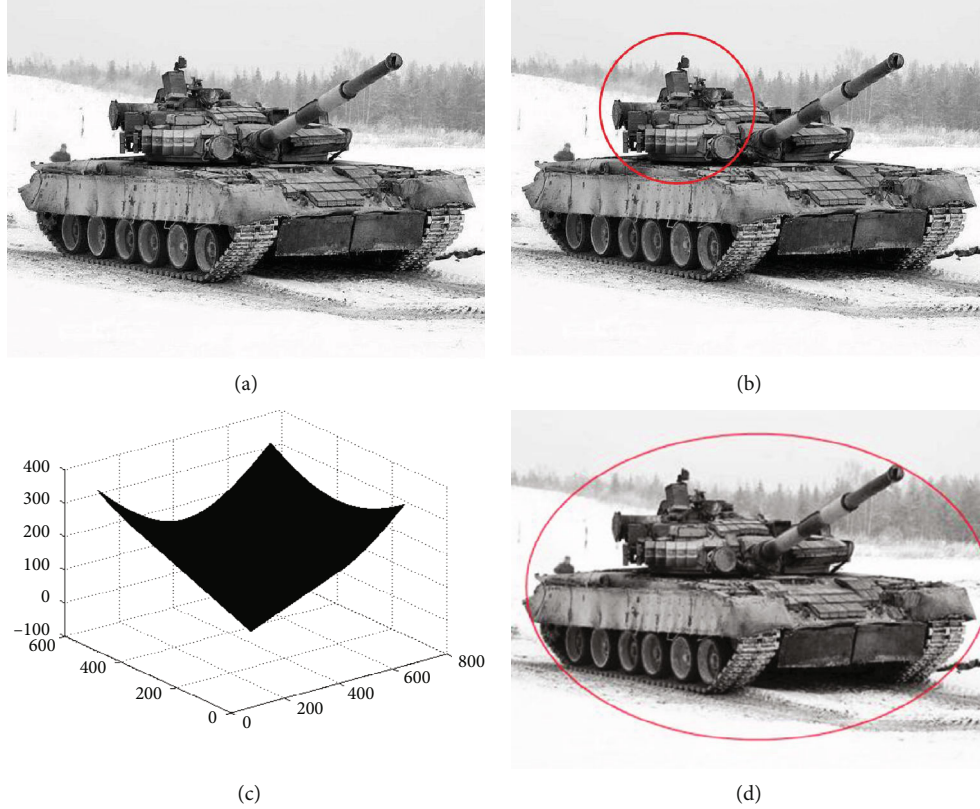


FIGURE 5: Accurate positioning results of T72.

Let $O-x-y-z$ be the coordinate system of the camera, where O is the optical origin, and a point $P(X, Y, Z)$ in space corresponds to a point $P'(X', Y')$ on the image plane, then:

$$\frac{Z}{f} = \frac{X}{X'} = \frac{Y}{Y'} \quad (10)$$

where f is the focal length of the camera.

Let $O-u-v$ be the pixel coordinate system, u and v axes are parallel to x axis to the right and y axis down, respectively. If the pixel coordinates are scaled α times on the u axis and β times on the v axis, the relationship between the coordinate P' and the pixel coordinates $[u, v]^T$ is:

$$\begin{cases} u = \alpha X' \\ v = \beta Y' \end{cases} \quad (11)$$

Let $\alpha f = f_x$, $\beta f = f_y$, be rewritten into a matrix form through a homogeneous linear equation as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \stackrel{def}{=} \frac{1}{Z} KP \quad (12)$$

where K is the parameter matrix in the camera.

By formula (12), the real coordinates of the target in space can be obtained, and then it can be combined with

the CapsNetv2 to realize the recognition and positioning of the radar target.

4. Algorithm Implementation

Figure 6 shows the specific process of the proposed radar target image recognition and positioning model based on the CapsNetv2.

- (1) Input target images such as BTR70, BMP2 and T72 in the MSTAR dataset as different output vectors into the CapsNetv2
- (2) The primary capsule layer is obtained through the convolution operation of the 1×1 reduction layer and the 3×3 convolution layer. Then the two image capsule layers are, respectively, trained and predicted to obtain $8 \times 8 \times 255$ and $16 \times 16 \times 255$ two prediction boxes
- (3) The dynamic routing algorithm iterative formula (1) is updated to obtain the characteristic capsule layer
- (4) According to formula (4), the maximum probability of the vector output modulus can be obtained, and the classification with the maximum probability of the radar target is obtained, and the two prediction boxes are mutually verified
- (5) A clear and complete edge line is obtained through edge extraction of the identified target classification image

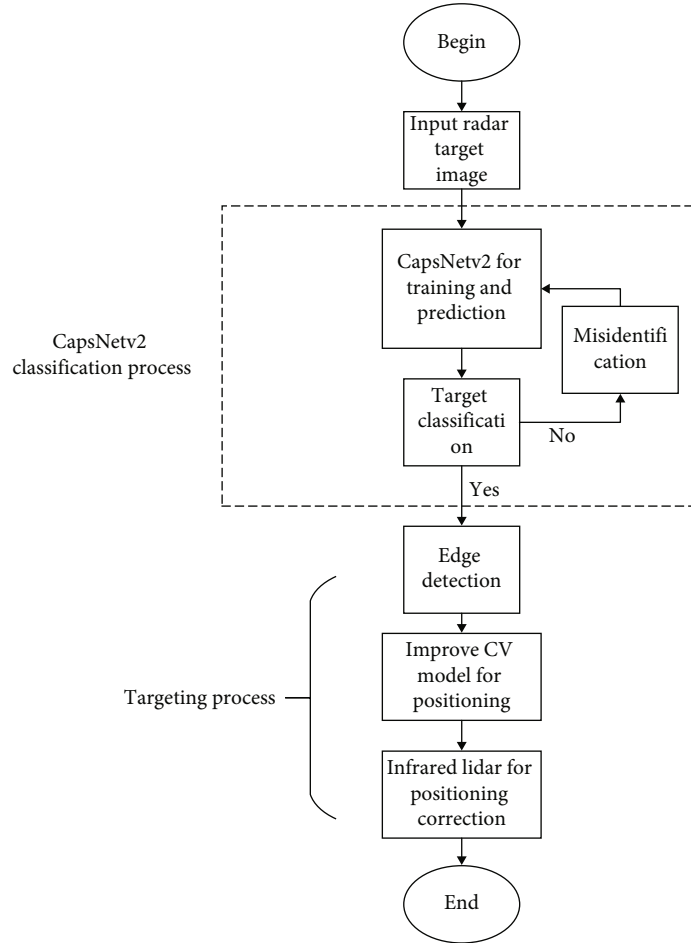


FIGURE 6: Flowchart of the proposed radar target recognition and positioning.

- (6) By improving the CV model and the precise positioning of the infrared lidar, the real three-dimensional coordinates of the target can be obtained

5. Experimental Verification

The method proposed in this paper is implemented using MATLAB R2014b and TensorFlow software. The 6000 radar target images with complex backgrounds in the MSTAR dataset are used as the training set, and 20% of the training set is randomly selected as the test set to verify the accuracy of the classification.

Example target images are shown in Figure 7. Among them, (a) and (b) are the T72 tanks in the sand and forest environments, (c) and (d) are the BTR70 armored vehicles in the sand and forest environments, and (e) and (f) are the BMP2 tank in the sand and forest environments.

Figure 7 shows that the CapsNetv2 can accurately identify radar targets in different complex backgrounds and has good robustness.

5.1. Different Network Training Effects. To verify the practicability and recognition accuracy of the CapsNetv2, the

radar target images with different complex backgrounds were used for training, and the performance of the CapsNetv2 was compared with that of the deep learning models of AlexNet and YOLOv4. The learning rate and the step length were changed and their performances were compared to select the best value of the parameter. The learning rate was 0.5, and the total number of steps was equal to 3000 as the optimal parameter. The training results are represented by the loss values, as shown in Figure 8.

Figure 8 shows the loss functions of AlexNet, YOLOv4 and CapsNetv2. The following conclusions can be drawn:

- (1) The Loss value has shown an overall downward trend in the training of the three networks, and the first half of the decline is very fast. However, compared with the AlexNet and YOLOv4 networks, the initial loss value of the CapsNetv2 is only 0.9. This is because the AlexNet needs to scramble every time it reads the data, while in YOLOv4, the MSE loss itself has certain problems and needs to be replaced by IOU loss
- (2) After the step size reaches 3000, the final loss function value of the CapsNetv2 is equal to 0.00015,



FIGURE 7: Identification effect.

which is about ten times smaller than the loss value of AlexNet. This is because the CapsNetv2 uses a simpler convolutional layer and a protocol layer, and two image capsule layers for training. At the same time, a model can be selected that has not been trained over-fitting, so that the model has good robustness

- (3) Since there are no corner feature points in the AlexNet, and training on a small sample dataset cannot make the model more stable, once the loss value reaches 0.5, the model training becomes jittered

and the training is terminated early. However, the CapsNetv2 retains different features information and training is more stable, which highlights the superior performance of the CapsNetv2

In order to verify and improve the learning performance of the capsule network, a database was used to randomly select the image data and compared with several other different algorithms. The results are shown in Figure 9.

It can be seen from Figure 9 that the recognition rate of the CapsNetv2 is higher than that of AlexNet and YOLOv4.

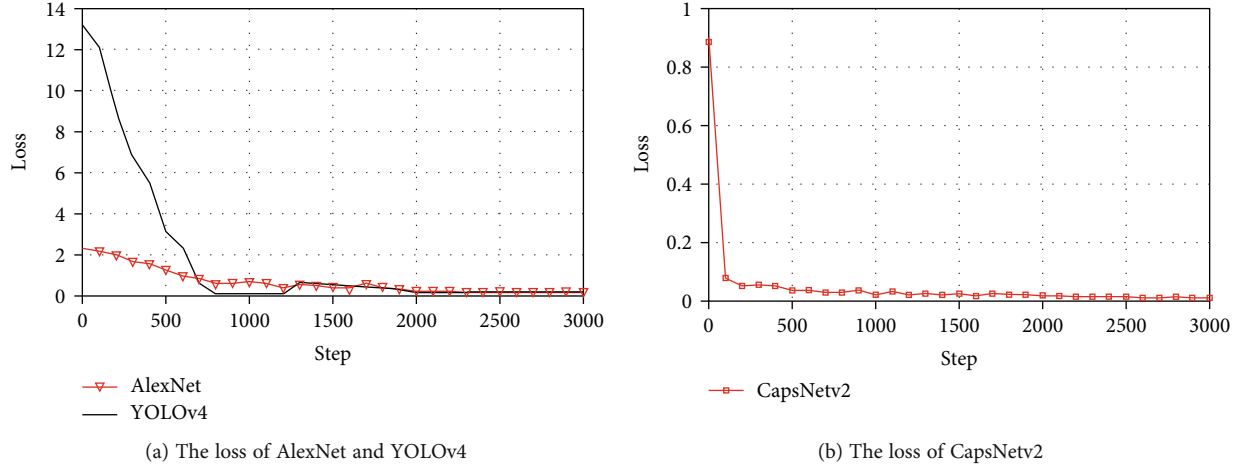


FIGURE 8: Loss values of different training models.

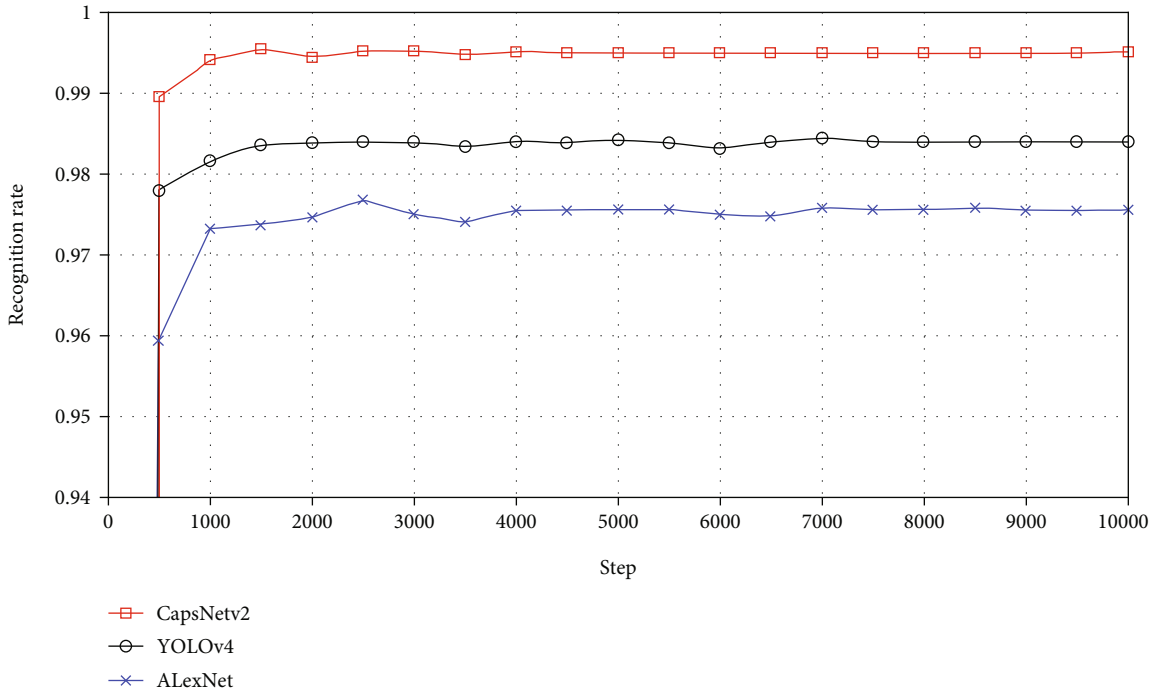


FIGURE 9: Contrasts of identification rates of different networks.

TABLE 2: Identification times of different algorithms.

	400*600	600*800	800*100
AlexNet	7.6	8.2	12.4
Yolov4	8.4	8.9	11.6
CapsNetv2	2.5	2.9	4.3

Through continuous learning, the recognition accuracy reaches 99.5%. This is the result of learning by multiple vector capsules and retaining different feature vectors (such as amplitude and angle). At the same time, the two image capsule layers can be predicted separately, which reduces the phenomenon of over-fitting and the possibility of misclassification.

Table 2 compares the recognition times of different algorithms. It can be seen from the table that compared with AlexNet and YOLOv4, the CapsNetv2 has a shorter classification time and is more suitable for detecting radar targets in different complex backgrounds.

5.2. Target Positioning in a Complex Background. The improved CV model is used to locate the radar target image identified in the CapsNetv2. From Steps 4 to 6 in Figure 6, the precise positioning of the radar target includes edge extraction, CV model positioning and infrared lidar positioning correction. The positioning results are shown in Figure 10.

It can be seen from Figure 10 that the improved CV model proposed in this paper and then corrected by the



FIGURE 10: The results of radar target positioning.

TABLE 3: Comparison of positioning accuracies of different methods.

Positioning method	Target precise positioning accuracy rate/%
The CV model+ infrared lidar positioning	97.5%
Infrared lidar imaging	92.6%
Local feature analysis	90.2%

infrared lidar can accurately locate the radar target images in different backgrounds.

From the different types of radar targets identified by the CapsNetv2, a group of 20 images were randomly selected for precise positioning of the radar target, and compared with the infrared imaging method and the local feature analysis

method [18]. Table 3 compares the positioning accuracies of different methods.

Table 3 shows that the proposed radar target recognition and positioning method that combines the CapsNetv2 and the CV model is more suitable for small sample learning and has better training effects. Thus, the proposed method has higher positioning accuracy than the other methods and is suitable for different complex backgrounds.

6. Conclusion

With the continuous improvement of military warfare technology, the real-time detection of different radar targets under different complex backgrounds is particularly important. This paper proposes a radar target detection model based on the CapsNetv2 and the improved CV model modified by infrared lidar. The proposed model can identify radar targets in complex backgrounds and accurately locate

their positions. The target positioning algorithm is simulated and experimentally verified. The following conclusions can be drawn:

- (1) The CapsNetv2 has strong self-learning and adaptive capabilities, and has a good training effect for small sample sets. It can effectively detect different types of radar targets and suppress interference caused by complex backgrounds. The recognition rate reaches as high as 99.5%. The reason is that the input of CapsNetv2 is a vector, which retains the feature information of the target to the greatest extent, and through the double-layer image capsule layer for training, it effectively reduces the over-fitting phenomenon and can more accurately classify different radar targets
- (2) The radar target image identified and classified by CapsNetv2 is segmented by the improved CV model, and finally corrected by the infrared laser mine, which can accurately locate the position of the target. The accuracy rate of the proposed method reaches 97.5%, which is more suitable for the precise positioning of radar targets than the other methods

The method proposed in this paper can better realize the radar target recognition under complex background and can provide accurate location information to meet the requirements of real-time inspection. However, the training time of the CapsNetv2 for a large number of images is relatively long. Thus, reducing the training time of the capsule network will be the focus of future research.

Data Availability

The [MSTAR] data and the [CapsNetv2 solution] data used to support the findings of this study were supplied by [Jiaxing Hao] under license and so cannot be made freely available. Requests for access to these data should be made to [Jiaxing Hao, 873328461@qq.com].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Y. Shen and X. Y. Huang, "Doppler weather radar wave echo identification and effect verification based on Bayesian classifier," *Marine Sciences*, vol. 44, no. 6, pp. 83–90, 2020.
- [2] J. R. Liu, "Weather radar rapid identification and early warning of heavy rainfall based on probability statistical model," *Information Recording Materials*, vol. 20, no. 4, pp. 157–158, 2019.
- [3] Z. Zhang, C. Zou, P. Han, and X. Lu, "A runway detection method based on classification using optimized Polarimetric features and HOG features for PolSAR Images," *IEEE Access*, vol. 8, pp. 49160–49168, 2020.
- [4] Z. Zhou, Q. M. J. Wu, S. Wan, W. Sun, and X. Sun, "Integrating SIFT and CNN feature matching for partial-duplicate image detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 5, pp. 593–604, 2020.
- [5] Y. Feng, "Mobile terminal video image fuzzy feature extraction simulation based on SURF virtual reality technology," *IEEE Access*, vol. 8, no. 99, pp. 156740–156751, 2020.
- [6] W. Q. Liu, Z. G. Liu, A. Núñez et al., "Multi-Objective Performance Evaluation of the Detection of Catenary Support Components Using DCNNs," *IFAC PapersOnLine*, vol. 51, no. 9, pp. 98–105, 2018.
- [7] S. Huang, Y. Zhai, M. Zhang, and X. Hou, "Arc detection and recognition in pantograph-catenary system based on convolutional neural network," *Information Sciences*, vol. 501, pp. 363–376, 2019.
- [8] Z. Li, L. Yang, and L. Wang, "Detection approach based on an improved faster RCNN for brace sleeve screws in high-speed railways," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4395–4403, 2019.
- [9] Y. Gan, H. Wu, and N. Xiao, "Cross-modal attentional context learning for RGB-D object detection," *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1591–1601, 2019.
- [10] Z. Zhou, Y. Cao, M. Wang, E. Fan, and Q. M. J. Wu, "Faster-RCNN based robust coverless information hiding system in cloud Environment," *IEEE Access*, vol. 7, pp. 179891–179897, 2019.
- [11] Y. Guo, Z. Pang, and J. Du, "An improved AlexNet for power edge transmission line anomaly detection," *IEEE Access*, vol. 47, no. 8, pp. 97830–97838, 2020.
- [12] S. Du, P. Zhang, and B. Zhang, "Weak and occluded vehicle detection in complex infrared environment based on improved YOLOv4," *IEEE Access*, vol. 9, pp. 25671–25680, 2021.
- [13] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic Routing Between Capsules," *Advances in neural information processing systems*, vol. 30, pp. 90–96, 2017.
- [14] J. Zh and X. Chen, "An improved overlapping handwritten digit recognition method for capsule networks," *Laser Journal*, vol. 40, no. 7, pp. 43–46, 2019.
- [15] C. Xiang, Z. Lu, and W. Zou, "MS-CapsNet: A Novel Multi-Scale Capsule Network," *IEEE Signal Processing Letters*, vol. 25, no. 12, pp. 1850–1854, 2018.
- [16] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion Recognition from Multiband EEG Signals Using CapsNet," *Sensors*, vol. 19, no. 9, p. 2212, 2019.
- [17] H. Y. Zhao, Z. J. Liu, and H. Zhang, "Multi-moving target segmentation based on improved CV model," *Chinese Journal of Scientific Instrument*, vol. 31, no. 5, pp. 1082–1089, 2010.
- [18] Q. F. Lai, J. Yang, and L. W. Han, "Insulator automatic identification and defect diagnosis model based on YOLOv2 network," *Chinese Power*, vol. 9, no. 22, pp. 1–10, 2019.