Non-Cooperative Target Recognition of Optical Remote Sensing Images based on Deep Learning Combined with Spatio-Temporal Reasoning

Xiaogang Tang¹*, Lin Zhu¹, Jun Peng²
¹Space Engineering University, Beijing, China
²Beijing spatial information transmission technology research Center, Beijing, China

*Corresponding author e-mail: titantxg@163.com

Abstract  In order to solve the low utilization rate of spatio-temporal sequence information during non-cooperative target recognition of optical remote sensing images with deep learning method, this paper puts forward a new method for offshore non-cooperative target recognition and tracing based on spatio-temporal reasoning. In the first stage, this method uses YOLOv3 model through deep learning to recognize characteristics of large batch of offshore non-cooperative target images, and finish preliminary screening of suspicious target; in the second stage, it uses time weight function and Markov model to process and analyze temporal and spatial sequence information respectively during continuous tracing of the target and get the probability distribution of certain specific target in the two dimensions; in the third stage, it uses D-S evidence theory to process probability information in temporal and spatial dimensions and get target probability with higher reliability through fusion reasoning. The experimental verification shows the comprehensive recognition precision in the first stage is over 92%; subsequently the secondary recognition precision of the target can be improved for 35% through fusion reasoning of time weighted position distribution probability and Markov position transfer probability. Results show the reasoning elements of spatio-temporal sequences obvious improve precision of secondary discovery after the non-cooperative target is lost, and provide new thinking for non-cooperative target tracing with intelligent method.

1. Description of non-cooperative target recognition problem

In recent years, with ceaseless improvement of ocean resource development, utilization and shipbuilding level, the maritime transportation becomes more busier, and non-cooperative targets are increasing in quantity. The accurate and rapid detection, recognition and tracing of non-cooperative targets by comprehensively using various types of information play a significant role in promoting economic development, maintaining maritime rights, fighting against crimes, carrying out offshore rescue and realizing accurate guidance. The rapid development of HD earth observation technology makes spatial resolution of remote sensing images higher, detailed characteristics of the target richer, and texture more complex; in the meantime, the data volume keeps increasing, and the target recognition is increasingly difficult. The offshore target recognition algorithm, which has high detection and recognition precision and low counting costs, becomes one of the focuses in current study.
The traditional target recognition technology extracts characteristics through artificial approaches by algorithms of HOG and HSV, and then uses SVM (Support Victor Machine) and other classifiers to realize image target recognition. However, those methods are restricted by artificial factors. The precision of artificial extraction of characteristics plays a crucial role in results of detection and recognition. With development of the Convolutional Neural Network (CNN) and maturity of deep learning theory [1], a series of mature target inspection and recognition algorithms, in which, the most representative ones are: Region-based CNN (R-CNN) [2], Spatial Pyramid Pooling Convolutional Networks (SPP-Net) [3], Fast Region-based CNN (Fast R-CNN) [4], Faster Region-based CNN (Faster R-CNN) [5], and other two-stage methods and some one-stage methods including YOLO (You Only Look Once) [6], Single Shot Multi-Box Detector (SSD) [7], etc..

Aforesaid algorithms provide new solutions for target extraction and inspection of offshore remote sensing images. However, compared to images in natural environment, remote sensing images have multiple factors such as diversified target dimension, special perspective, complex background, diversified direction and cloud interference. The problem to be concerned and solved is to apply mature and efficient algorithm to remote sensing image processing field. Document [8] studies and proves that the target inspection and recognition method based on Faster R-CNN could rapidly and accurately recognize remote sensing image, and could be well popularized. Document [9] proposes labeling method of rotating rectangular framework aiming at difficult characteristic extraction under complex background, and designs rotating target recognition with RR-CNN. Document [10] combines image sampling and characteristic pyramid network to improve recalling rate and precision of small size targets and meanwhile verify the possibility to move target inspection algorithm of natural picture through deep learning to remote sensing images for target recognition. Document [11] reduces complexity of calculation while guaranteeing inspection precision by improving YOLO model.

Aforesaid methods mainly study image dimensions deeply, and have favorable recognition effect for targets. However, large amount of spatio-temporal sequence information has not been utilized sufficiently. Meanwhile, aforesaid method is inadequate in continuous tracing and supervision of offshore non-cooperative target. Particularly, it is difficult to establish again the recognition and tracing process after the target is lost.

2. New method based on deep learning combined with spatio-temporal reasoning

Thus, this paper, based on multi-class offshore target inspection according to aforesaid algorithm, further studies extraction and utilization of spatio-temporal sequence behind image information, and expands tracing and recognition of non-cooperative target from single image to multiple dimensions of images, time and space. By sufficiently utilizing multi-source and multi-dimensional information for target recognition and fusion seasoning, it improves precision of offshore non-cooperative target discovery, recognition and tracing.

The overall framework of the method can be seen in Figure 1. In Stage 1, YOLOv3 model, time weight function and Markov model are used to extract characteristics and collect sequences from three dimensions of images, time and space, and generate non-cooperative target image recognition mode, time weight probability matrix and Markov position transfer matrix; for the characteristic that traces of non-cooperative target are difficult to predict, and target is easy to lose during tracing, when the target is lost, YOLOv3 recognition model is used firstly in Stage 2 to check images of the suspicious region, and find multiple suspicious targets in different areas; meanwhile confirm time weight probability distribution of the target before losing in different areas and the Markov position transfer probability from place before lost to other places in accordance with time weight matrix and Markov position transfer matrix; in Stage 3, the final probability distribution of target place is acquired through information fusion of D-S theory in accordance with two probability distributions acquired in the previous stage.
3. Specific implementation process of the method

3.1. YOLOv3 model multi-class target inspection based on improved anchor box cluster method

YOLO model is a target recognition method based on regression, and could realize end-to-end training with small counting costs and fast calculation speed. It more complies with rapid processing demand of massive remote sensing information in current and a future stage. This paper improves YOLOv3 model.

3.1.1. K-means cluster method improvement of distance adjustment formula. In the aspect of anchor, YOLOv3 network introduces anchor boxes in Faster R-CNN\cite{12}, and VOC and COCO data sets are both predicted from three dimensions. In the meantime, the algorithm adopts K-means algorithm to cluster the data set according to size of target framework so as to avoid subjective influence caused by manual setting of transcendence frame, in favor of subsequent training and prediction. The experiment re-clustered data set. This paper wishes to acquire IOU relevant to box size through anchor boxes. Therefore, K-means algorithm often uses measurement ways such as Euclidean distance, Manhattan distance, Chebyshev, which are inapplicable to the application background of this paper\cite{14}.

This paper improves the distance formula as follows:

\[ d(\text{box}, \text{centroid}) = 1 - IOU(\text{box}, \text{centroid}) \] (1)

In the formula, \( \text{centroid} \) means the center of the cluster, \( \text{box} \) refers to sample, \( IOU(\text{box}, \text{centroid}) \) refers to Intersection over Union (IOU) of cluster center frame and K-means frame, and IOU means precision degree of the prediction frame. The formula is set below:

\[ IOU(b_{y0}, b_{w0}) = \frac{b_{y0} \cap b_{w0}}{b_{y0} \cup b_{w0}} \] (2)

3.1.2. YOLOv3 model characteristic extraction and loss function. YOLOv3 algorithm uses Darknet53 to extract image characteristics. Meanwhile, when using Darknet-53, the full connection layer is removed, and \( 1 \times 1 \) convolution kernel is used as output so as to guarantee input of images with any size. It only constitute characteristics extraction network by convolutional layer character extraction and residual layer control and training effect. The loss function of the algorithm is as follows:

\[ Loss = \text{bbox loss} + \text{confidence loss} + \text{class loss} \] (3)

The loss function mainly optimizes three outputs of YOLOv3 model: the prediction frame size, degree of confidence and class. The detailed explanation of the three factors are as follows:

(1) Loss of prediction frame size:

\[ \text{bbox loss} = \sum_{0}^{1} \text{obj} \ast [ (b_{x} - l_{x})^2 + (b_{y} - l_{y})^2 + (b_{w} - l_{w})^2 + (b_{h} - l_{h})^2 ] \] (4)
In the formula, $1^{obj}$ means the target, $(b_x,b_y)$ and $(l_x,l_y)$ are respectively predicted and actual central coordinate, and $(b_x,b_y)$ and $(l_x,l_y)$ are respectively predicted and actual width and height.

(2) Confidence loss:

$$
\text{confidence loss} = \sum_{0}^{n} KL(p_{0}, q_{0})
$$

In the formula, $p_{0}$ is the detected confidence, and $q_{0}$ is actual value. The cross entropy of the two factors will be calculated.

(3) Class loss:

$$
\text{class loss} = \sum_{0}^{n} 1^{obj} \sum_{c=0}^{C} KL(p(c), q(c))
$$

In the formula, $1^{obj}$ means the target, $C$ is total class, $C=3$ in this paper, and $p(c), q(c)$ are respectively probability of each class and actual probability; the cross entropy of the two factors will be calculated.

3.2. Construction of probability distribution matrix based on time weight function

In the data acquired, we select aircraft carrier as target class for test. For images with unclear hull number, the canny edge detection algorithm is used to process details of images to acquire hull number information; the target data of the class is segmented according to null number (every carrier has the only null number) and shooting place to establish target set $\{N_i\}_{i=1}^{N}$ and geological information port set $\rho(T) \geq 0, \forall T \geq 0$.

3.2.1. Confirmation of time weight function

During mathematical statistics of time sequence, in order to make berthing probability of certain object in different ports more comply with actual situations. The time weight function is adopted to weight the timing of image acquisition. The time weight function shall meet the monotonic increasing within the entire definitional domain. The smaller the corresponding function value on the data point approaching to the origin is, the larger the function value on the data set far from the origin will be. Moreover, the difference between two adjacent data points shall not be too large, and the growth trend of function shall be smooth.

(1) $\rho(T) \geq 0, \forall T \geq 0$, and in monotonic increasing;
(2) Weight increase $\Delta \rho(T) = \rho(T) - \rho(T - 1)$ is small, and the function is growing smoothly;
(3) $\lim_{T \to \infty} \rho(T) = 1$.

During sort-out of time sequence data, Jan.1, 2010 is selected as time benchmark $t_0$, and day as the unit is used to calculate difference between image shooting time and benchmark time $T_n = t_n - t_0$, $T_n \in (0,3600)$, and shall be positive integer. The function image created is shown in the figure. The nature of the function complies with aforesaid conditions.

$$
\rho(T) = \frac{1}{2} \left[ \frac{2}{\pi} \arctan \left( \frac{T}{180} \right) + 1 \right]
$$
3.2.2. Construct time weight function. In K port, the carrier with null number of \( N_j \) acquires weight of \( \rho_{S,N_j}(T_n) \) at the time of \( T_n \).

Total weight that the carrier acquires in the port shall be:

\[
\sum_{k} \rho_{S,N_j}(T_n) 
\]  (8)

The sum of weight that the carrier acquires in different ports shall be:

\[
\sum_{k} \sum_{n} \rho_{S,N_j}(T_n) 
\]  (9)

The weighted probability distribution of the same carrier in different places through normalization shall be:

\[
P_{(N_j,S_k)} = \frac{\sum_{k} \rho_{S,N_j}(T_n)}{\sum_{k} \sum_{n} \rho_{S,N_j}(T_n)} 
\]  (10)

The weighted probability distribution matrix is built as follows:

\[
T_{(N_j,S_k)} = \begin{bmatrix}
P_{(N_j,S_k)} & P_{(N_j,S_2)} & \cdots & P_{(N_j,S_k)} \\
p_{(N_j,S_k)} & P_{(N_j,S_2)} & \cdots & P_{(N_j,S_k)} \\
\vdots & \vdots & \ddots & \vdots \\
p_{(N_j,S_k)} & P_{(N_j,S_2)} & \cdots & P_{(N_j,S_k)} 
\end{bmatrix} 
\]  (11)

In the matrix, \( P_{(N_j,S_k)} \) means the weighted time probability of carrier \( N_j \) when berthing at \( S_k \) port, and \( \sum_{k=1}^{N} P_{(N_j,S_k)} = 1 \).

3.3. Construction and improvement of Markov position transfer matrix based on statistical method

According to the spatio-temporal sequence information of different carrier berthing points, the statistical method is used to construct Markov position transfer matrix. In one time period, one carrier \( N_j \) is observed at long-term equal intervals. The movement between different places will generate a
transfer status set. The target transfers between different statuses. The frequency of transfer between
different statuses is collected to respectively get transfer status matrix \( X \) and status frequency matrix \( n \).

3.3.1. Markov position transfer matrix based on statistical method. The matrix \( X \) contains all
transfer statuses of carrier \( N_i \); \( X_{ab} \) means the status of the carrier transferring from port \( a \) to port \( b \),
and \( n_{ab} \) means frequency of such status in the entire observation period:

\[
X = \begin{bmatrix}
X_{11} & X_{12} & \cdots & X_{1b} \\
X_{21} & X_{22} & \cdots & X_{2b} \\
\vdots & \vdots & \ddots & \vdots \\
X_{a1} & X_{a2} & \cdots & X_{ab} \\
\end{bmatrix}
\]

\[
n = \begin{bmatrix}
n_{11} & n_{12} & \cdots & n_{1b} \\
n_{21} & n_{22} & \cdots & n_{2b} \\
\vdots & \vdots & \ddots & \vdots \\
n_{a1} & n_{a2} & \cdots & n_{ab} \\
\end{bmatrix}
\]

The Markov position transfer matrix for the carrier is shown as below:

\[
M_{N_i} = \begin{bmatrix}
m_{11} & m_{12} & \cdots & m_{1b} \\
m_{21} & m_{22} & \cdots & m_{2b} \\
\vdots & \vdots & \ddots & \vdots \\
m_{a1} & m_{a2} & \cdots & m_{ab} \\
\end{bmatrix}
\]

\[
m_{ab} = \frac{n_{ab}X_{ab}}{\sum_{b=1}^{b=n} n_{ab}X_{ab}}
\]

(12)

In the formula, \( m_{ab} \) means the Markov position transfer probability of certain carrier from port \( a \) to
port \( b \), and \( \sum_{b=1}^{b=n} m_{ab} = 1 \).

3.3.2. Correction of position transfer matrix. According to aforesaid algorithm, the probability of
certain carrier for every transfer status can be acquired so as to infer the probability distribution of the
carrier in next time point. However, for errors generated by uncertainty and randomness of transfer
status statistics, the following two statuses may exist: Firstly, certain less frequent transfer status
incurred actually but was not collected actually, leading to zero probability distribution of the status
probability, which is seriously inconsistent with actual situations; secondly, certain carrier may
transfer in the port observed; it may not leave the port actually but be judged by mistake to transfer to
other places due to manmade factors. Therefore, in order to weaken aforesaid two interference factors
to certain degree, it is necessary to add rectification coefficient \( \psi \) to aforesaid transfer matrix. In case
of aforesaid extreme circumstances, the probability distribution will be revised according to the
following rules.

\[
M(m_{ab})=\begin{bmatrix}
m_{a1} & m_{a2} & \cdots & m_{ab} \\
\end{bmatrix}, \forall m_{ab} \neq 0
\]

\[
M(m_{ab})=\begin{bmatrix}
\frac{m_{a1}-\psi}{\sum_{b=1}^{b=n} m_{ab}-\psi} & \frac{m_{a2}-\psi}{\sum_{b=1}^{b=n} m_{ab}-\psi} & \cdots & \frac{m_{ab}-\psi}{\sum_{b=1}^{b=n} m_{ab}-\psi} \\
\end{bmatrix}, \forall m_{ab} = 0
\]

(13)
3.4. Spatio-temporal information integrated reasoning and recognition based on D-S evidence theory

The time weight ideas believe the shipping and berthing of carriers have certain rules, and the recent activities have larger weight since they could reflect various situations in current period. Markov model believes from a long-time scale, the shipping and berthing of carriers may use one status transfer matrix to describe the probability distribution of carriers berthing in certain port in future certain time transferring to other ports. Aforesaid two models are difficult to describe the actual situations. For non-cooperative target lost during tracing, after acquiring a batch of targets from suspicious area through image processing, in order to precisely position targets lost before, the method proposed in this paper confirms a probability distribution from aforesaid time weighted probability distribution matrix and Markov position transfer matrix, and improve precision of judgment through spatio-temporal information integrated reasoning.

Suppose the target to be traced is \( N_i \), the position is \( S_n \) port at the \( t \) time before the tracing target is lost; according to time weight probability distribution matrix, the time weight probability distribution in different ports shall be:

\[
T_{N_i|S_n} = \begin{bmatrix} P_{(N_i|S_1)} & P_{(N_i|S_2)} & \cdots & P_{(N_i|S_n)} \end{bmatrix}
\]

(14)

The position transfer probability distribution from \( S_n \) port to other ports shall be:

\[
M(m_{S_n}) = \begin{bmatrix} m_{S_1, S_n} & m_{S_2, S_n} & \cdots & m_{S_n, S_n} \end{bmatrix}
\]

(15)

From the realistic perspective, aforesaid two probability distributions are difficult to describe the shipping and berthing rules of carriers due to restriction of respective model generation; meanwhile, the moving rules of offshore non-cooperative targets are uncertain and random to a great degree, the conflict between probability distribution acquired by multiple modes will change continuously. Therefore, the integrated reasoning model shall be stable. For different conflict degrees, favorable information integration effects shall be provided. D-S evidence theory has complete mathematical foundation, and is a common method for intelligent information processing. This paper applies D-S evidence theory to integrated reasoning of spatio-temporal information.

D-S evidence theory is used to integrate aforesaid two probability distribution to get new probability distribution \( P_{D-S} = [S_1 \ S_2 \ \cdots \ S_n] \), and the integration rules are as follows:

\[
F(\beta) = K^{-1} \sum_{s_i \cap m_k \neq \emptyset} T(s_i)M(m_{s_i})
\]

(16)

\[
K = \sum_{s_i \cap m_k \neq \emptyset} T(s_i)M(m_{s_i})
\]

In the formula, \( K \) is orthogonalization coefficient,

The target after integrated reasoning shall meet the condition \( P_{D-S}(S_{\text{target}}) = \max \{P_{D-S}(S_i)\} \).

4. Experiment and results

The method in this paper contains three stages including image space mode training, spatio-temporal information processing and judgment.

4.1. Image space mode training stage

This paper establishes offshore target data set through multiple channels. The images are 1476 historical remotely sensed images from GoogleEarth historical remotely sensed data, China resources satellite and Changguang No.1 satellites; the pixels are about 750pixel×750pixel and 1200pixel×1200pixel. During the experiment, due to limited number of images collected, the sample number is insufficient. In order to guarantee quality of model trained, and prevent from unobvious extraction characteristics and over-fitting during training, 50% images selected from data set are applied with image enhancement technology at random such as color jitter, random cropping, random
rotation and noise interference \cite{14} (one image may be applied with several image enhancement technologies), as shown in the figure.

![Diagram of image enhancement](image)

**Figure 3.** Diagram of image enhancement

The data set after image enhancement contains 2692 pieces different types of images for targets with different sizes, which are divided into three classes including warcraft, merchant ship and aircraft carriers. The labeling software is used to mark the images and made to data set in the format of VOC2007. There are 1252 warcraft samples, 1417 merchant ship samples, and 869 aircraft carriers (one image may contain several samples).

In the experiment, firstly, the reconstructed K-means clustering algorithm is used to test according to different K values, and different Average IOU results can be seen in the Figure. According to experiment results, Average IOU is increased with the rising of K value. When K value is around 7, the increase rate slows down. Thus, the number of optimal anchor box is 7. The initial rectangular frame acquired by K-means algorithm is respectively: [40, 64], [93, 35], [161, 676], [316, 471], [561, 88], [584, 74], [665, 100].

![K-means clustering results](image)

**Figure 4.** K-means clustering results

In the experiment, the output sizes of YOLOv3 network after adjustment are respectively 13×13×24, 26×26×24 and 52×52×24. YOLOv3 network structure sets that every grid unit predicts 3 boxes, and each box shall have five basic parameters (x, y, w, h, confidence). Meanwhile, the probability of aforesaid three classes shall be predicted. The three depths outputted finally by the model are 24.
The model apply initialization with the weight parameter acquired by training based on PASCAL VOC2007 data set. The initial learning rate is 0.001, and linear attenuation is adopted.

**Table 1.** Experiment configuration and parameters

| System configuration | Specific parameters |
|----------------------|---------------------|
| Operating system     | Ubuntu16.04         |
| GPU                  | NVIDIA GeForce RTX 2060 |
| Memory               | 32GB                |
| Deep learning framework | Keras               |

![Image of result analysis](image)

**Figure 5.** Visualized diagram of loss value.

The model acquired by training have higher recognition probability for targets of merchant ship, warcraft and aircraft carrier. The experiment results can be seen in Figure 6.

![Image of experiment results](image)

**Figure 6.** YOLOv3 image space experiment results.
Table 2. Objective evaluation index of the model

| Target type     | Average recognition precision P/% |
|-----------------|----------------------------------|
| Merchant ship   | 95.14                            |
| Warcraft        | 92.69                            |
| Aircraft carrier| 90.83                            |
| average         | 92.89                            |

4.2. Spatio-temporal information joint treatment

4.2.1. Probability distribution matrix based on time weight function. Aiming at spatio-temporal sequence data collected, the time weight method is used to time weight probability of six ships $N_1, ..., N_6$ to calculate; the time weight probability distribution matrix is as follows:

$$
P = \begin{bmatrix}
21.79\% & 59.58\% & 18.63\% \\
30.02\% & 43.28\% & 26.70\% \\
19.68\% & 5.36\% & 74.96\% \\
11.55\% & 36.31\% & 52.14\% \\
82.71\% & 14.92\% & 2.37\% \\
1.53\% & 33.12\% & 65.35\%
\end{bmatrix}
$$

4.2.2. Markov position transfer matrix based on statistical method. Take the ship $N_1$ as an example. According to its berthing rules, it is confirmed to have the following transfer statuses:

$$A \rightarrow B; B \rightarrow A; A \rightarrow C; B \rightarrow C; C \rightarrow B; C \rightarrow A$$

The transfer matrix and status frequency matrix of aforesaid transfer status are respectively:

$$X = \begin{bmatrix}
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}$$

$$n = \begin{bmatrix}
0 & 5 & 7 \\
4 & 0 & 3 \\
2 & 5 & 0
\end{bmatrix}$$

According to aforesaid matrix, the Markov position transfer matrix shall be:

$$M_{x_1} = \begin{bmatrix}
0 & 41.67\% & 58.33\% \\
57.14\% & 0 & 42.86\% \\
28.57\% & 71.43\% & 0
\end{bmatrix}$$

Aforesaid probability may generate extreme circumstance of zero probability. Therefore, according to the correction method, $\psi=0.05$ is used to make correction:

$$M_{x_1} = \begin{bmatrix}
5.26\% & 38.63\% & 56.11\% \\
54.88\% & 5.26\% & 39.86\% \\
24.81\% & 69.93\% & 5.26\%
\end{bmatrix}$$
4.3. Information integration based on D-S evidence theory

For target tracing problem of \( N \), aforesaid data shows when the carrier followed is lost at Port C, the probability distribution of the carrier probability in port A, B and C according to time weight matrix shall be

\[
T_{(N\mid C)} = \begin{bmatrix}
0.2179 & 0.5958 & 0.1863 \\
\end{bmatrix}
\]

According to the Markov position transfer matrix, three transfer statuses may take place: it may stay in somewhere not discovered in Port A, or it is ahead toward Port B and Port C. The probability distribution of aforesaid three statuses shall be

\[
M(C) = \begin{bmatrix}
0.2481 & 0.6993 & 0.0526 \\
\end{bmatrix}
\]

The results of information integration based on D-S evidence theory shall be

\[
P_{D-S} = \begin{bmatrix}
0.1125 & 0.8671 & 0.0204 \\
\end{bmatrix}
\]

In accordance with results, the carrier only transferred to port B. The precision of judgment is increased 45.54% and 24% respectively compared to the simple time weight method or Markov method.

Aforesaid method is used to apply spatio-temporal integrated seasoning for the transfer process of six carriers \( N_1, \ldots, N_6 \). The experiment results can be seen in Table 3:

| Carriers | Seasoning method | Probability distribution | Judgment results \( S_{target} \) |
|----------|-----------------|--------------------------|-------------------------------|
| \( N_1 \) | \( T_{(N\mid C)} \) | 0.2179 0.5958 0.1863 | B |
| \( N_2 \) | \( M(C) \) | 0.2481 0.6993 0.0526 | B |
| \( N_3 \) | \( P_{D-S} \) | 0.1125 0.8671 0.0204 | |
| \( N_4 \) | \( T_{(N\mid B)} \) | 0.3171 0.2214 0.4615 | C |
| \( N_5 \) | \( M(B) \) | 0.3171 0.2214 0.4615 | C |
| \( N_6 \) | \( P_{D-S} \) | 0.0965 0.0326 0.8709 | |

5. Conclusion

This paper puts forward a new method for non-cooperative target recognition based on spatio-temporal sequence seasoning for the problem of difficult tracing after target is lost. Firstly, it uses YOLOv3 target detection model to find out a batch of suspicious targets from the image dimension through deep learning, and then respectively builds transfer probability distribution from the dimensions of time and space respectively based on time and space acquired during continuous monitoring of non-cooperative targets. Next, it applies D-S evidence theory to integrated seasoning of aforesaid two probability distributions so as to improve precision of offshore non-cooperative target
judgment. In terms of verification method, this method could make comprehensive judgment by integrating temporal and spatial information so as to improve precision and success rate of re-tracing after offshore non-cooperative target is lost. The experiment results show compared to the traditional artificial interpretation which is affected greatly by manmade and subject factors, this method proposes quantified model and standard to provide auxiliary decision making to artificial interpretation. This method improves precision and efficiency of artificial interpretation by screening large probability targets and eliminating small probability targets.

References

[1] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.

[3] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.

[4] GIRSHICK R. Fast R-CNN[C] // Proceedings of 2015 IEEE International Conference on Computer Vision. Chilen: IEEE, 2015: 1440-1448.

[5] REN S Q, HE K M, GIRSHICK R S, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C] // Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779-788.

[7] LIU W, ANGUELOV D, R HAN D, et al. SSD: single shot multi-box detector[C] // Proceedings of 2016 European Conference on Computer Vision. Netherlands: Springer, 2016: 21-37.

[8] Wang J C, Tan X C, Wang Z H, et al. Faster R-CNN deep learning network based object recognition of remote sensing image[J]. Journal of Geo-information Science, 2018, 20(10): 1500-1508. DOI: 10.12082/dqxxkx.2018.180237

[9] Zikun Liu, Jingao Hu, Lubin Weng, Yiping Yang: Rotated region based CNN for ship detection. ICIP 2017: 900-904.

[10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of 2014 IEEE Conference on computer vision and pattern recognition. Columbus, OH, USA: IEEE, 2014: 580-587.

[11] MA Xiao, SHAO Limin, JIN Xin, et al. Improved YOLO model and its application in ship target recognition[J]. Telecommunication Engineering, 2019, 59(8): 869-874.

[12] REDMON J, FA R HADI A. YOLO V3: an incremental improvement[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2018: 1-6.

[13] ZHENG Z Q, LIU Y Y, PAN C C, et al. Application of improved YOLO V3 in aircraft recognition of remote sensing images[J]. Electronics Optics & Control, 2019, 26(4): 28-32.

[14] Zhou M, Shi Z W, Ding H P. Aircraft classification in remote-sensing images using convolutional neural networks Journal of Image and Graphics, 2017, 22(5): 702-708.