

Research and Application of Target Detection Algorithm

Hui Qiu

School of Merchant Shipping, Shanghai Maritime University, Shanghai, 201306, China

Abstract

Target tracking is a difficulty in the field of computer vision. By tracking the target, vehicle and ship in the video data, we can extract the moving track of the target, and then analyze the behavior law of individual or group. Video target tracking has become an important research content of computer vision, and plays a more and more important role in public security control, human-computer interaction, traffic control, military and other fields. Although scholars at home and abroad have proposed a variety of target tracking algorithms and built a relatively perfect target tracking system, the robustness of the algorithm is still a big challenge. The deep tracking algorithm is introduced in detail from the perspective of deep target detection. The application of deep learning in video target detection algorithm is described, and its future development is prospected.

Keywords

Object tracking; Computer vision; Trajectory extraction.

1. Introduction

In recent years, the wave of deep learning swept the field of computer vision, which not only improved the detection performance of general objects, but also greatly promoted the development of target detection, and laid a good foundation for the research in the field of target tracking [1]. Target tracking is different from target detection. It is necessary to use data association technology to associate the target with the largest similarity between the two frames in order to achieve the purpose of continuous tracking of the target in the video [2] [3], so as to obtain the speed, trajectory and direction of the target motion [4], and further apply it to the research field of individual or large-scale group behavior [5] [6]. Object tracking is a basic task in the application of computer vision.

With the wide application of deep learning technology in the field of computer vision, it has become the mainstream of academia to use deep learning method to study object tracking [6-8]. In order to make up for the lack of existing literature, and make the majority of scientific research workers grasp the latest development trend in the field of target tracking, this paper introduces the algorithm in the field of target tracking, according to the classification method of traditional tracking algorithm and deep learning tracking algorithm [9, 10], and then selects the corresponding index to evaluate the performance.

2. Traditional Tracking Algorithm

2.1. Kalman Filtering Algorithm

In 1960, in order to solve the problem of linear filtering of discrete data, Kalman filter algorithm was proposed. The algorithm was later extended to the field of target tracking [11]. Kalman [12] prediction is an algorithm that estimates the minimum variance of the tracking target's motion trajectory, so as to predict the next frame position of the moving target. Through the establishment of the observation equation of the moving target and the actual motion equation of the moving target, the two equations are used for Kalman prediction to obtain the next frame position estimation of the moving target [13]. The advantage of Kalman prediction algorithm is

its small amount of calculation and good real-time performance of the algorithm, but the algorithm is only suitable for the motion system conforming to Gaussian distribution [14].

The innovation of the algorithm is that the errors in the prediction process and the measurement process are considered at the same time, and the two kinds of errors exist independently and are not affected by the measurement data [15]. Kalman filter algorithm is often combined with other algorithms to complete the target tracking task, rarely used alone.

2.2. Particle Filter Algorithm

Aiming at the shortcoming that Kalman filter requires the state variables of the target to satisfy Gaussian distribution, Xiang [16] proposed a multi-target tracking and detection algorithm based on particle filter framework, which is a general method of Kalman filter algorithm. Particle filter is suitable for nonlinear, non Gaussian system, it is developed from Kalman filter. Particle filter obtains the similarity distribution of each particle position image and the target by sampling particles near the position of the tracking target in the last frame, and then samples the importance of the particle position, which forms the real-time tracking of the target [17]. The calculation of particle filter is less, so the algorithm has good real-time performance and is widely used in the tracking algorithm.

Because the salient features of the algorithm come from many underlying characteristics of the target, the algorithm has strong robustness. At the same time, particle filter can achieve the optimal estimation of nonlinear non Gaussian state space model [18]. Therefore, the algorithm can deal with the problem of track management in the process of multiple target tracking, as well as the problems of target appearance, disappearance, merging, splitting and being blocked by obstacles.

2.3. Mean Shift Algorithm

Mean shift algorithm is another algorithm to reduce the search range. Mean shift is a nonparametric estimation method. Cheng [19] calculates the similarity between the template in the previous frame and the candidate template in the current frame, and selects the template with the highest similarity to get the mean shift vector of the target. Mean shift algorithm is a nonparametric statistical method to search the pattern closest to the distribution of sample points. Mean shift algorithm is an effective method in color image sequence target tracking. However, in the infrared target tracking, due to the lack of information to describe the infrared target in a single gray feature space, the infrared target tracking based on mean shift algorithm is not robust.

3. Deep Learning Algorithm

3.1. Convolutional Neural Network

Convolutional neural network (CNN) is a typical deep learning model. Lecun [20] proposed the concept of CNN. In 2012, Krizhevsky Applied DCNN to the field of image classification for the first time, and successfully introduced DCNN into the field of computer vision. According to the characteristics of infrared oversampling scanning imaging, Lin [21] proposed an infrared point target detection method based on deep convolution neural network. The basic structure of CNN includes input layer, convolution layer, pooling layer, full connection layer and output layer, see Figure 1.

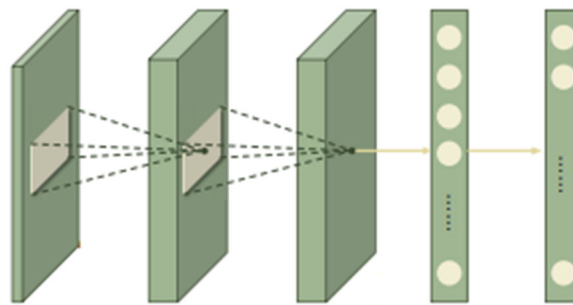


Figure 1. Two or more references

In the actual network, there are often several convolution layers and pooling layers connected alternately. In the convolution layer, the filter is usually used to perform convolution operation to extract the features of the image. The shallow convolution layer in front of the network is used to extract the low-level features of the image, and the deeper convolution layer is used to extract the high-level features of the image [22]. The full connection layer normalizes the extracted features of the image into one-dimensional feature vectors, and the output layer is used to output the general information of each category in the classification problem Rate.

3.2. RCNN Network Model

In 2014, Chen [23] proposed RCNN model for target detection, and introduced convolutional neural network into the field of target detection. RCNN uses CNN method to extract features in the target candidate region, see Figure 2.

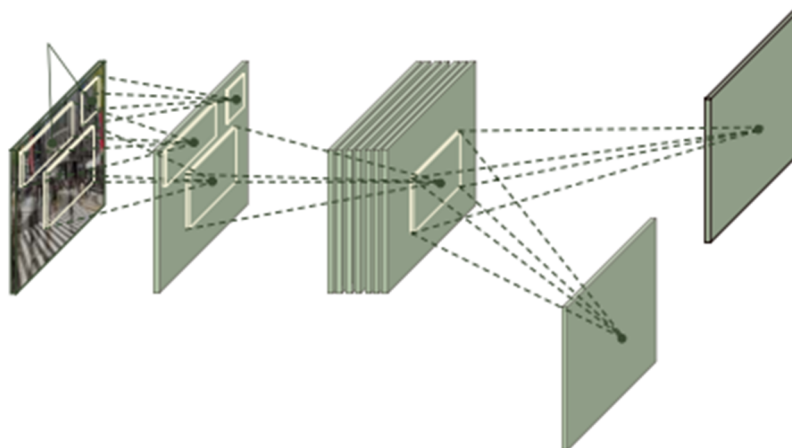


Figure 2. Flow chart of RCNN algorithm

The core of SPP net network model is to add a spatial pyramid pooling layer after the convolution layer, which can generate a fixed size image without clipping the image, thus reducing the loss of features. Moreover, in the whole process, only one convolution feature extraction is performed on the image, which greatly improves the target detection speed of general objects [24].

3.3. YOLO Series Model

The target detection algorithms described above all divide the target candidate regions first, and then predict the target categories. However, YOLO treats the division of target candidate regions and the prediction of target categories as a regression problem, and outputs the positions and categories of multiple targets directly on the image, so as to realize the rapid detection of targets on the premise of ensuring high accuracy. YOLO uses the improved

perception V1 model to extract the features of the image [25], takes the adjusted whole image as the input of the convolution network, and divides the original image. At this time, the center point of the object in the image will fall in a grid cell, and the corresponding grid cell is responsible for detecting the object. Each grid cell predicts B candidate boxes and the confidence score within the candidate box.

Finally, the category prediction information and confidence score are encoded into the vector as the output of the YOLO output layer. In the prediction, each grid cell will generate a candidate box according to the category probability, but large objects will generate multiple candidate boxes. The author uses the non maximum suppression algorithm to select the candidate box with the highest score of intersection and union ratio, and remove the redundant window to optimize the detection results.

3.4. Multi Target Tracking Evaluation Index

Because the average accuracy of the evaluation index of deep learning algorithm belongs to the target detection field, it can only be used to measure the accuracy of the detected target, but can not be used to measure the tracking performance of multi-target algorithm. Therefore, in order to compare the performance of the above tracking algorithms, it is necessary to select the corresponding multi-target tracking evaluation index to measure it. In reference, two evaluation indexes, Mota and motp, were first proposed, see Table 1.

Table 1. Performance comparison of different algorithms

Algorithm	MOTA↑	MOTP↑	MT↑	ML↓	IDS↓
Kalman filtering	85.00%	-	-	-	-
Particle filter	-	-	80.80%	0.70%	10
Multi target tracking	30.30%	71.30%	13.00%	38.40%	680
YOLOv3	60.50%	79.30%	30.20%	19.60%	1129

4. Conclusion

The concept of machine learning and other algorithms is introduced into the field of ship image recognition, some more advanced image data, and synthetic aperture radar system (SAR) is also introduced into the field of ship image recognition. Through the image recognition of the SAR, the feature values of the ship's aerial top view are extracted, and the results obtained by KNN classifier, Bayesian classifier and neural network classifier are combined with the ship features as the input of SVM, and the more accurate classification results are obtained. The experimental results show that the method can extract and recognize the characteristics of SAR images Other effects are better. The computer can recognize the ship type from the satellite perspective, which increases the technology of automatic identification system to practical application.

Target tracking is a difficult and hot issue in the field of computer vision, and it is also a difficult problem in human behavior dynamics. There is still a long way to go for the combination of computer vision based target tracking technology and artificial intelligence, and more efforts need to be put into the research of recognition algorithm. Literatures have used deep learning technology to analyze and apply detection target tracking. This makes people in the continuous development of new algorithms for target tracking analysis provides a new idea.

References

- [1] YANG Min, PEI MingTao, WU YuWei, JIA YunDe. Learning online structural appearance model for robust object tracking [J]. Science China, 2015, 58(03): 105-118.

- [2] Gwang-Min CHOE, Tian-jiang WANG, Fang LIU, Chun-Hwa CHOE, Hyo-Son SO, Chol-Ung PAK. An advanced integrated framework for moving object tracking [J]. Journal of Zhejiang University-Science C(Computers & Electronics), 2014, 15(10): 861-877.
- [3] Effective method for tracking multiple objects in real-time visual surveillance systems [J]. Journal of Systems Engineering and Electronics, 2009, 20(06): 1167-1178.
- [4] Micha Livne, Leonid Sigal, Nikolaus F. Troje, David J. Fleet. Human attributes from 3D pose tracking[J]. Computer Vision and Image Understanding, 2012, 116(5).
- [5] Jenhwa Guo. Optimal measurement strategies for target tracking by a biomimetic underwater vehicle[J]. Pergamon, 2008, 35(5-6).
- [6] R. Hayashi, N. Yanagisawa. Characteristics of the visco-elasticity of the human muscle during visual-guided tracking with the hand[J]. Electroencephalography and Clinical Neurophysiology, 1995, 95(4).
- [7] Yi Huang, Yingmin Jia. Adaptive fixed-time relative position tracking and attitude synchronization control for non-cooperative target spacecraft fly-around mission[J]. Journal of the Franklin Institute, 2017, 354(18).
- [8] Yang Juqing, Wang Dayong, Zhou Weihua. Precision laser tracking servo control system for moving target position measurement[J]. Optik - International Journal for Light and Electron Optics, 2017, 131.
- [9] Qing E. Wu, Tuo Wang, Yong Xuan Huang, Ji Sheng Li. State fusion of fuzzy automata with application on target tracking[J]. Computers and Mathematics with Applications, 2008, 57(6).
- [10] Hongyang Xue, Yao Liu, Deng Cai, Xiaofei He. Tracking people in RGBD videos using deep learning and motion clues[J]. Neurocomputing, 2016, 204.
- [11] Hongkai Yu, Dazhou Guo, Zhipeng Yan, Lan Fu, Jeff Simmons, Craig P. Przybyla, Song Wang. Weakly supervised easy-to-hard learning for object detection in image sequences[J]. Neuro computing, 2020, 398.
- [12] Chensi Cao, Feng Liu, Hai Tan, Deshou Song, Wenjie Shu, Weizhong Li, Yiming Zhou, Xiaochen Bo, Zhi Xie. Deep Learning and Its Applications in Biomedicine [J]. Genomics, Proteomics & Bioinformatics, 2018, 16(01): 17-32.
- [13] Jie Ren, Zulin Wang. A Novel Deep Learning Method for Application Identification in Wireless Network [J]. China Telecom, 2018, 15(10): 73-83.
- [14] LIN Liangkui, WANG Shaoyou, TANG Zhongxing. Using deep learning to detect small targets in infrared oversampling images[J]. Journal of Systems Engineering and Electronics, 2018, 29(05): 947-952.
- [15] Huafeng Chen, Jun Chen, Ruimin Hu, Chen Chen, Zhongyuan Wang. Action Recognition with Temporal Scale-Invariant Deep Learning Framework[J]. China Telecom, 2017, 14(02): 163-172.
- [16] Yikun Yang, Shengjie Jiao, Jinrong He, Bisheng Xia, Jiabo Li, Ru Xiao. Image retrieval via learning content-based deep quality model towards big data [J]. Future Generation Computer Systems, 2020, 112.
- [17] Anil Kumar, C.P. Gandhi, Yuqing Zhou, Rajesh Kumar, Jiawei Xiang. Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images[J]. Applied Acoustics, 2020, 167.
- [18] Ao Ding, Yuan Zhang, Lei Zhu, Yanping Du, Luping Ma. Recognition method research on rough handling of express parcels based on acceleration features and CNN[J]. Measurement, 2020, 163.
- [19] Haidar Almubarak, Yakoub Bazi, Naif Alajlan. Two-Stage Mask-RCNN Approach for Detecting and Segmenting the Optic Nerve Head, Optic Disc, and Optic Cup in Fundus Images[J]. Applied Sciences, 2020, 10(11).
- [20] Yuwei Lu, Yuan Yuan, Qi Wang. A dense connection based network for real-time object tracking[J]. Neurocomputing, 2020, 410.
- [21] Rodrigo Bazo, Eduardo Reis, Lucas Adams Seewald, Vinicius Facco Rodrigues, Cristiano André da Costa, Luiz Gonzaga, Rodolfo Stoffel Antunes, Rodrigo da Rosa Righi, Andreas Maier, Björn Eskofier,

- Rebecca Fahrig, Tim Horz. Baptizo: A sensor fusion based model for tracking the identity of human poses[J]. Information Fusion, 2020, 62.
- [22] Hua Lan, Jirong Ma, Zengfu Wang, Quan Pan, Xiong Xu. A message passing approach for multiple maneuvering target tracking [J]. Signal Processing, 2020, 174.
- [23] Jiakai Liao, Guoliang Xiang, Libo Cao, JiaHao Xia, Luyao Yue. The left-behind human detection and tracking system based on vision with multi-model fusion and microwave radar inside the bus[J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2020, 234(9).
- [24] Megumi Ota, Hiroshige Tateuchi, Takaya Hashiguchi, Takehiro Kato, Yasuhiro Ogino, Momoko Yamagata, Noriaki Ichihashi. Verification of reliability and validity of motion analysis systems during bilateral squat using human pose tracking algorithm[J]. Gait & Posture, 2020, 80.
- [25] Yang Su, Ting Cheng, Zishu He, Xi Li. Adaptive simultaneous multibeam resource management for colocated MIMO radar in multiple targets tracking [J]. Signal Processing, 2020, 172.