Image recognition Technology in Agricultural Vertical Field based on CNN and pre-training Model

Xu Bailing[1], Zhang Wanli[2], Ma Xiaoyu[3]

[1]College of Mathematics and Computer Science, Northwest University for Nationalities, Lanzhou, Gansu 730000

[2]Chengdu College of University of Electronic Science and technology of China, Chengdu, Sichuan 610000

[3]College of Electronic Information Engineering, Hebei University, Baoding, Hebei, 071000

Keywords: artificial intelligence; target recognition; convolution neural network; feature fusion; transfer learning

Abstract: In order to realize the accurate identification of cow individuals in the complex farm environment, the SSD algorithm is improved to solve the problem that the (single shot multibox detector) algorithm is not effective in detecting overlapping objects. First of all, the feature fusion of different feature images can make different feature images complement each other and improve the detection effect of overlapping objects; then, remove the Conv4_3 layer from the network and increase the number of candidate boxes of other feature images, which can not only ensure the real-time performance of the algorithm but also improve the detection accuracy; finally, the transfer learning method is introduced to improve the average accuracy of the algorithm. The experimental results show that: compared with the traditional SSD algorithm, the improved SSD algorithm improves the average accuracy AP (average precision) by 4.32% while satisfying the real-time detection; after migration, the improved SSD algorithm AP increases by 3.85%.

1. Introduction

With the rapid development of domestic dairy industry, large-scale farming has become the mainstream direction. Individual identification of dairy cows is very important to the management of dairy cows in pasture [1-2]. The recognition of dairy cows has mainly experienced "artificial observation method" [3], "ear cutting method" [4], "electronic recognition method" [5] and so on. These methods will not only cause damage to dairy cows, but also tedious work and high labor intensity.

In order to realize the accurate recognition of dairy cows, the SSD algorithm [9] is improved in this paper. First of all, in order to make full use of the feature information of each feature layer, the feature image of SSD algorithm is fused, and the shallow feature map containing detail information is fused with the deep feature map containing semantic information, which is not only beneficial to target classification, but also more accurate to predict the position of objects. Then, combined with the characteristic that there are almost no small target objects in the FriesianCattle2017 data set, the Conv4_3 layer in the network is removed. At the same time, the number of other feature map candidate boxes is increased to improve the detection accuracy of the algorithm. Finally, transfer learning is introduced.

2. Improved convolution neural network

2.1 Change SD to accumulate meridians and collaterals

In the SSD network structure, the shallow feature map has higher resolution and contains more detailed information, but because of less convolution, it has lower semantics and more noise. The deep feature map has stronger semantic information, but the resolution is very low and the ability to perceive details is poor. In order to make full use of shallow detail information and deep semantic information, this paper combines deep feature map with shallow feature map to enhance the detection effect of overlapping objects.

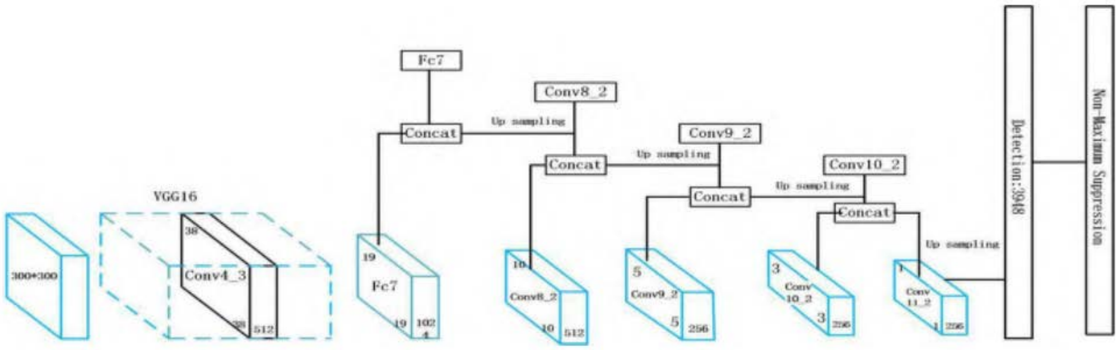

Figure 1 Improve the structure of SSD network

As shown in figure 1, the network structure of the SSD algorithm is improved. The backbone network of the improved SSD algorithm is VGG16,. Compared with the SSD algorithm, this algorithm has five output feature images, and the five feature images are fused by the method of channel merging to enhance the detection effect of overlapping objects. At the same time, the candidate boxes are reconstructed, so that there are 3948 candidate boxes in the improved SSD algorithm.

2.2 Feature fusion mode

There are two kinds of feature fusion methods in the network model, ResNet, FPN and other network structures use element-wiseadd for feature fusion, while DenseNet and Inception use concat for feature fusion. As can be seen from figure 3, concat is the splicing of the number of channels, while add is the addition of feature graphs. Compared with add and concat, add shares one channel, so when merging, the number of channels of the two feature graphs must be the same. This results in fewer data dimensions and loss of information, and the weight of the additive part in feature fusion is always the same. Concat allows fusion information to have different parameters. Concat fusion increases the number of channels, and the network will extract more overall information of the object, but the amount of computation will also be larger. In order to choose which fusion method to use, this paper carries on the experiment under two different fusion methods.

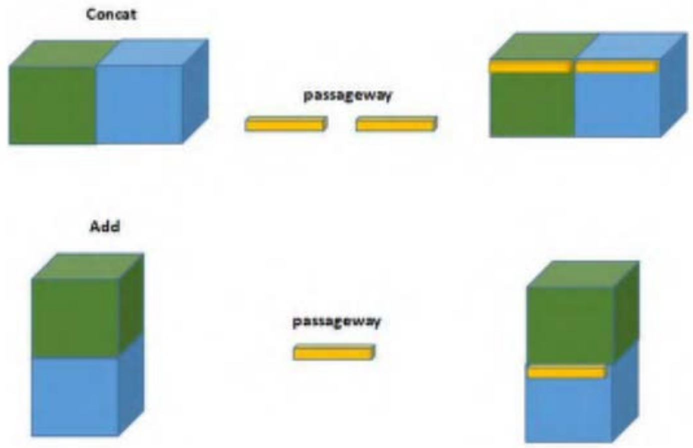

Figure 2 Different ways of fusion

3. Transfer learning

Transfer learning is a new strategy for small samples to supervise learning. Through the sharing of model parameters, the network models pre-trained in large-scale data sets are embedded into other task models as feature extractors for other tasks. As shown in figure 4, the model is trained from the source domain, and the knowledge learned from the source domain and the convolution neural network are transferred to the target domain. Then in the target domain, a newly designed classification layer and the migrated network model are reconstructed into a new convolution neural network, which is used to train the image data of the target domain. The key to transfer learning is the correlation between the source task domain and the target task domain. As shown in figure 3, the target task of this paper is to identify dairy cows. Then we can choose other related mammals as the original data domain to pre-train the model, and finally fine-tune the model to achieve a better recognition effect. In the experiment, two methods are used for training: first, the frozen feature extraction layer only trains the classification layer, and second, all layers are trained.

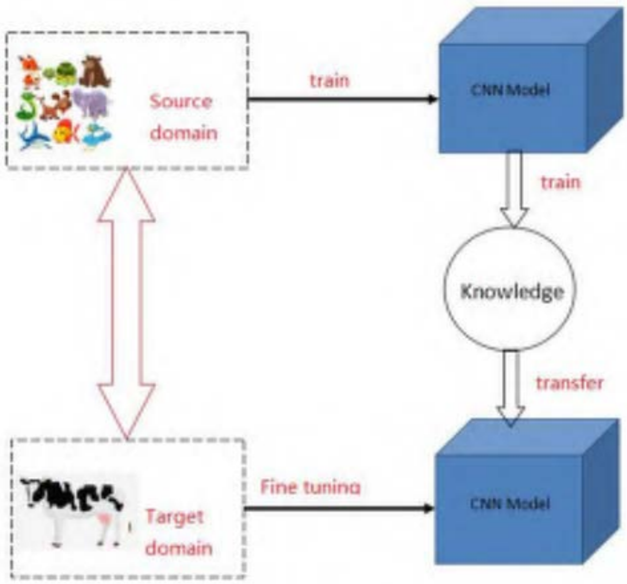

Figure 3 Transfer learning process

4. Experimental results and analysis

4.1 Experimental Environment

This experiment is carried out under the environment of python3.7, tensorflow2.0. The dataset is marked with labeling, the network structure is constructed with keras, and the training is accelerated with RTX2080Ti, CPU: i7-97003.6k@3.60GHZ*8.

4.2 Dataset

Because this paper uses less data for training, and the transfer learning method is highly targeted, which can well solve the over-fitting problem caused by the lack of training data, this paper uses the transfer learning method to train the model. In the experiment, the pre-training model was obtained by pre-training with VOC2012 data set. The cow data is a total of 940 FriesianCattle2017 data sets, which are divided into training set, verification set and test set according to the proportion of 6:2:2.

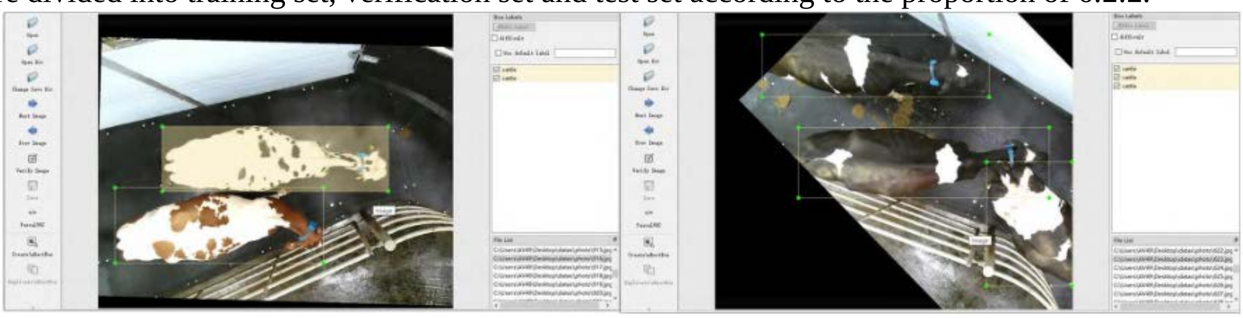

Figure 4 Cow label picture

4.3 Experimental results

As can be seen from Table 4, the average accuracy of the feature increase fusion method is the lowest. Although the average detection time of the channel splicing fusion method is long, it basically meets the requirements of real-time detection, and the average accuracy is 2.35% higher than that of the traditional method. 2.47% higher than the feature increase fusion method. Therefore, this paper chooses the way of channel splicing and fusion.

Table 1 Experimental results on test sets with different fusion methods

Experimental algorithm	Feature fusion	Fusion method	AP/%	Average detection tima/ms
SSD			88.23	46.23
SSD	√	add	88.11	51.52
SSD	√	concat	90.58	54.04

As can be seen from Table 1, although the average detection time of the improved SSD algorithm is increased by 7.93 Ms compared with the traditional SSD algorithm, the accuracy is increased by 1.13%, the recall rate is increased by 4.28%, the average accuracy is improved by 4.32%, and the performance of the algorithm is significantly improved.

Table 2 Experimental results of SSD algorithm and improved SSD algorithm on test set

Experimental algorithm	TP/FP/FN	Precision/%	Recall/%	AP/%	Average detection tima/ms
SSD	454/12/60	97.42	88.32	88.23	46.23
Improvement SSD	476/7/38	98.55	92.60	92.55	54.16

As can be seen from Table 2, although the average detection time of the improved SSD algorithm is increased by 7.93 Ms compared with the traditional SSD algorithm, the accuracy is increased by 1.13%, the recall rate is increased by 4.28%, the average accuracy is improved by 4.32%, and the performance of the algorithm is significantly improved.

In figure 5, the vertical line is used as the distinction, the leftmost is the SSD algorithm PR curve, and the far right is the improved SSD algorithm PR curve. Among them, the vertical line part is not in the PR curve, just to distinguish different PR curves and reflect the differences between different algorithms.

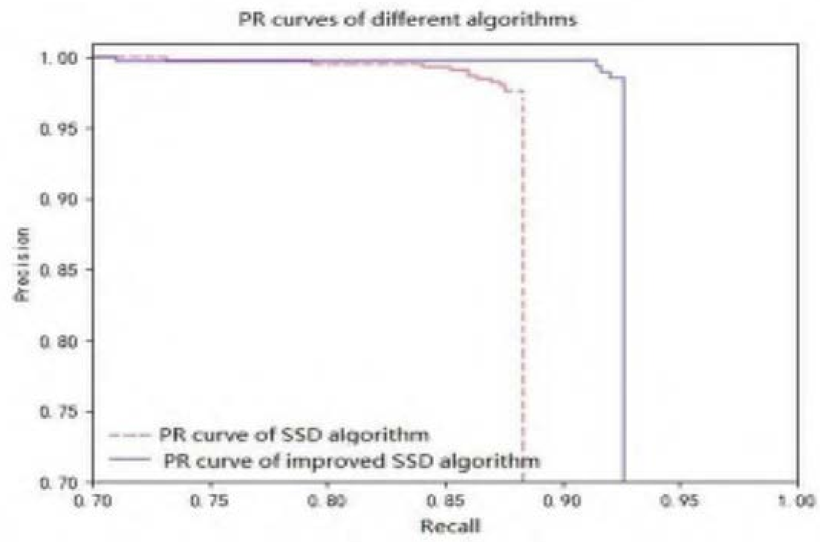

Figure 5 PR Curves of different algorithms

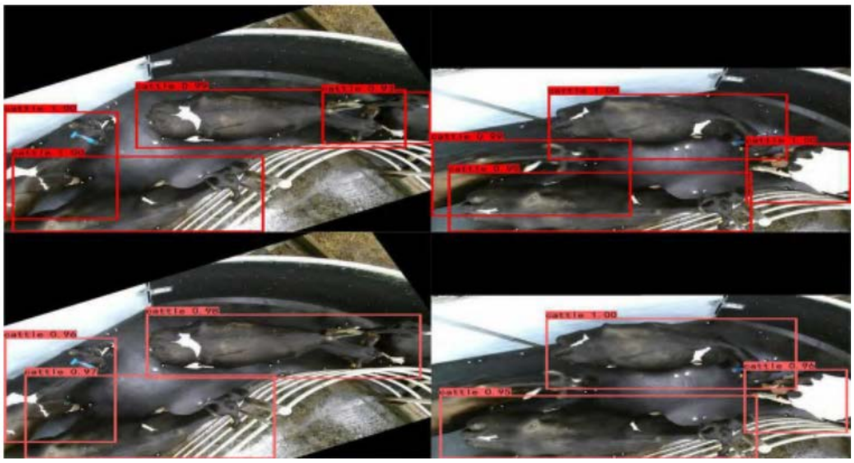

Figure 6 Detection effect of improved SSD algorithm and SSD algorithm

As shown in figure 6, in order to improve the detection results of the SSD algorithm and the SSD algorithm, the upper row is the detection results of the improved SSD algorithm, and the lower row is the detection results of the SSD algorithm. In the two images on the left, there are two objects with high coincidence and occlusion, so the SSD algorithm only detects three targets, while the improved SSD algorithm detects four targets; in the left two images, two objects overlap, so the SSD algorithm detects only three targets, while the improved SSD algorithm detects four targets. From the comparison of the detection results, it can be seen that the improved SSD algorithm improves the poor detection of overlapping objects by the SSD algorithm.

5. Conclusion

In order to solve the problem that SSD algorithm is not effective in detecting cow overlapping objects in FriesianCattle2017 data sets, an improved SSD algorithm is proposed in this paper. The feature fusion of different feature images is carried out, and the transfer learning training network is introduced to improve the detection accuracy of the algorithm. The experimental results show that the improved SSD algorithm can significantly improve the detection effect of dairy cows, and solves the problem of poor detection effect of SSD algorithm for overlapping objects. Compared with other mainstream algorithms, this algorithm has a good accuracy. It has a good application prospect in intelligent aquaculture.

Reference

[1] Ball J E, Anderson D T, Chan C S. A comprehensive survey of deep learning in remote sensing: theories, tools and challenges for the community[J]. Journal of Applied Remote Sensing,2017,11(4):042609.

[2] AyandaMS, DanaIS, AniatiMA,MohamadIF. CattleRaceClassificationUsingGrayLevel Co-occurrence Matrix Convolutional Neural Networks[J]. Procedia Computer Science,2015,59.

[3] Suresh N J, Satish K T, Sheng T, Huang D K. Recent advancement in biosensors technology for animal and live stock health management[J]. Biosensors and Bioelectronics,2017,98.

[4] Ali I A. From classical methods to animal biometrics: A review on cattle identification and tracking[J]. Computers and Electronics in Agriculture,2016,123.

[5] DhakadSK, DwivediU, BaudhaS, etal.PerformanceImprovementofFractalAntennawith Electromagnetic Band Gap (EBG) and Defected Ground Structure for Wireless Communication[M]//Gnanagurunathan G, Sangeetha R, Kiran K. Optical and Microwave Technologies.LectureNotesinElectricalEngineering,Singapore:Springer,2018,468:9-19.

[6] Xiong S H, Sun Z H. Design of cow yield measurement system based on RFID and image processing〔J〕. Computer application and software,2018,35(7):267-272.