Research on ship target recognition system based on Neural Network

Yan Wang
Hohhot Vocational College, Hohhot Inner Mongolia, 010070

Absrtact: Ship Target Recognition usually uses band image recognition technology, and multi band image fusion recognition can expand the application scope of recognition system. In order to improve the efficiency and accuracy of fusion recognition, this paper attempts to introduce convolutional neural network technology to design an intelligent fusion recognition method , which uses alexnet network model, from the visible light, medium wave, medium waveThe features of the three bands are extracted from the long wave infrared three band image, and the features of the three bands are screened by using mutual information to determine the feature vector with fixed length; According to the different feature levels, three fusion methods, early, middle and late, are used to verify the algorithm. The experimental results show that the accuracy of the method is the highest, reaching 84.5%.

Keywords: Target recognition;Ship identification; Feature fusion; Convolutional neural network; Multi band image; Feature selection

1 Research framework

1. Network architecture

To explore using CNN network model, we should first sort out multi band image data and resources. Here, we have built more than 5000 images, which are divided into six categories of targets; The three band image recognition method involved in this paper is mainly divided into three stages. Firstly, the parallel features of three band images are obtained through alexnet network; Secondly, the feature selection method based on mutual information is used to reduce the dimension of the obtained feature vectors and remove the irrelevant vectors; Finally, single quantity features are extracted and fused at different network layers (early fusion, middle layer fusion and late fusion).

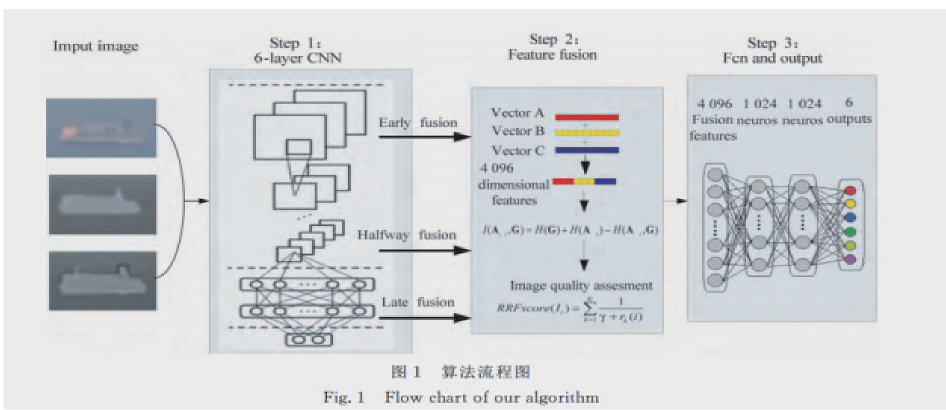

图 1 算法流程图
Fig. 1 Flow chart of our algorithm

Figure 1algorithm flow chart

In this paper, alexnet network model is used as the basis, and its parameters are modified. In the model initialization part, the training model in ilsvrc12 is used as the initial parameter by using the transfer learning characteristics of neural network, and then it is adjusted by using the multi band data established in this paper. In this study, three identical parallel networks are used to extract three different three band image features. C is the convolution layer, R is the nonlinear activation function (relu), l is the local response normalization, P is the maximum pool layer, f is the fully connected layer, and D is the regular function. Due to the different size of the image, we describe it as a 227×227 pixel rectangle to accommodate network input. The pixels in the input image are averaged and subtracted, and then trained by CNN network, and then extracted hierarchically according to the propagation of similarity. At the sixth level, 4096 dimensional feature vectors are obtained, which are called vector a. Similarly, in the medium wave and long wave infrared images, the same CNN network is used for feature extraction to obtain vecorb and vectorc respectively. The three sets of feature vectors contain the feature information of ship targets in different bands .

2. Mutual information feature selection based on adaptive weight

After series fusion, the dimension of the feature vector of the three immoral images is relatively high. The feature vector is adaptively selected through mutual information, and then the dimension of the feature vector is reduced .

The correlation between two variables is analyzed by mutual information. The mutual information of discrete random variables X and Y is expressed as follows:

$$I (x, y) = \sum_{y\in y} ^ {} \sum_{x\in x} ^ {} P (x, y) \log (\frac{p (x, y)} {p (x) P (y)}),$$

Where p (x) and P (y) represent the marginal probability distribution functions of X and y, and P (x, y) represents the joint probability distribution function.

And because

I (x; y) = H (x) + H (y) - H (x; y), where H (x) and H (y) are edge entropy and H (x, y) is joint entropy

Mutual information calculation between dimensions and labels:

Taking the visible image eigenvector vectora as an example, the i-dimensional vector of all image eigenvectors A_1 Put them together, and use G to express the vector, and use I for mutual information (A_1, g). Generally speaking, mutual information represents the importance of each dimension. The greater the mutual information, the higher the effectiveness of classification according to this dimension. Its calculation method is as follows:

I (a_ {: i}) =h (g) + H (a_ {: i}) - H (a_ {: i}, g)

Where, the image label G is a fixed value, and for different dimensions I, it remains unchanged, a: represents the number of samples, H represents the entropy of random variables, and the entropy of G also remains unchanged, so it only needs to be calculated

H (a_{: i}) -h9a_{: i}, g)

The size of mutual information can be sorted, and then the dimension of n-dimensional vector can be reduced as needed.

Concatenate features $F3_{CNN}$ Set to 4096 dimension, the threshold value is evaluated according to the image definition evaluation standard, and the comprehensive score of various evaluation indexes is calculated according to the extreme formula of RRF according to RRF method:

$$\mathrm{RRF}_{score}(I_i)=\sum_{k=1}^{K}\frac{1}{\gamma+r_{k(i)}}$$

Select the following five evaluation indicators:

The gradient similarity of pixels between images is calculated by similarity deviation;

Visual information fidelity index the image distortion degree is calculated based on multi-scale Gaussian mixture model of wavelet transform;

Calculate the similarity index of visible light;

Thirdly, the fidelity of the object structure in the image is calculated exponentially.

$r_{k(i)}$ Representation image I_i Ranking in the k-th evaluation criteria, γIs a constant, takeγ=60. Application formula $\mathrm{RRF}_{score}(I_i)=\sum_{k=1}^{K}\frac{1}{\gamma+r_{k(i)}}$ Analyze and sort the contribution rate of data in the data set. Here, we calculate through three band image RRF_{score} It is necessary to (I_i) And normalize it to obtain the weight values of eigenvectors of different bands. The calculation formula is as follows:

F-3cnn=\override{rrf (I \vis})} vectora + \override{rrf (I \mwir})} vectorb + \override{rrf (I \lwir})} vectorc

Wherein, {refers to normalization operation. After concatenation, the dimension of the three band image feature f-3cnn reaches 4096. Through normalization operation, the band images with poor definition are screened out, and the dimension of the feature vector is reduced, so as to reduce the error. By normalizing the feature vectors after series fusion through the full connection layer and the output layer, the classification probabilities for different targets are obtained. Each layer of the full connection layer contains 1024 neurons, and the output layer uses the softmax function to classify different types of ship targets.

3. Feature fusion at different levels

This paper studies the fusion recognition at different levels, and inputs parallel networks at different levels and different bands. This paper is divided into three stages: early stage, middle stage and late stage.

Early fusion: after the first convolution layer, the feature vectors of visible and infrared images are fused in series, and then the features are selected through mutual information method to reduce the dimension of the feature vector after series fusion. The conv 1 layer in CNN network extracts the low-level visual features of the image, that is, the detail texture information, including edges, corners, points and so on. It can be seen that the fusion model in the early stage is aimed at the low-level features of the image.

Middle layer fusion: in the convolution layer, the feature vectors of different band images are concatenated after the conv 4 layer, and the feature is also selected through mutual information method to reduce the dimension of the feature vector. The semantic content of conv 4 convolution layer is more detailed than that of conv 1 convolution layer, and the detailed texture information is relatively clear.

Later fusion: the feature vectors of the full connection layer FC 6 are fused in series, and the dimension is reduced by mutual information method to extract more accurate data. The results show that the information presented by the later fusion is more advanced and detailed.

2 Data set construction and training

The data set established in this paper has a resolution of 1024x768. The working frequency band of the MW detector is 3.7 ~ 4.8 μ m, the image resolution of 320x256, the working frequency band of the 8-14 μ m long wave sensor, and the imaging resolution of 640x480. The dataset includes 6 categories of targets and 5187 images. They are the cruise ship a354x3, the cruise ship b337x3, the railway ferry 3X3, the cargo ship 236x3, the 291x3 small fishing boat, and the 303x3 warship. The training set, validation set and test set were classified according to the order of 50%, 20% and 30% by random sampling. Using the (SGD) method, the initial learning rate of batch size M = 32, impulse 0.9, weight delay 0.0005 and 0.0.01, and the learning rate decreased to 0.001 and the learning cycle was 100 after the contemporary price function stabilized. The simulation verification platform is ubuntu140, i5-4590, gtx1080 video card, 16 g memory, Caffe deep learning architecture, which is used to build and train the network. If it is reused 100000 times, it will take about 4 hours.

3 Experimental verification and comparison

The target recognition rates of different fusion methods are tested through experiments, and the four methods are compared, and the causes of false recognition are analyzed.

(1) Hog+ support vector machine identification: select 64x128 image blocks with 8 steps, a total of 3780 dimensional vectors, and use hog feature descriptor for identification;

(2) SIFT feature recognition: 64x64 image blocks are selected and segmented into 128 dimensional SIFT features;

(3) Alexnet mode: initialize the selected parameters in ilsvrc12 and fine tune them using the data provided in this article;

(4) Vgg-16 mode: using the method similar to alexnet, the test layer is finely adjusted to improve the target identification ability . Because infrared imaging has the characteristics of low resolution, the details and textures it presents are less obvious than those of visible light, and the recognition accuracy of "bag of words model" alone is lower than that of manual extraction. The neural network recognition technology based on alexnet and vgg-16 has achieved satisfactory results in large-scale recognition, but because the model is only applicable to a single target, it can not give full play to the fusion characteristics between multispectral images, so it is difficult to achieve high-precision recognition under low resolution conditions .

The fusion recognition mode established in this paper makes full use of the complementary information of three band images, and uses the method of feature selection to eliminate the irrelevant vectors that have little effect on classific ation . The results show that the recognition rate of either fusion mode is higher than that of a single band. The results show that among the three methods, the recognition rate of image fusion using middle layer is 84.5%, 81.5%, 79.9% and 7%, respectively, and the recognition effect is the best. Through the analysis of the reasons, it is found that the low-level features extracted from the image, such as edges, corners, etc., are similar to the characteristics of manual design, close to the edges of manual design features, and there are deficiencies in the description model.

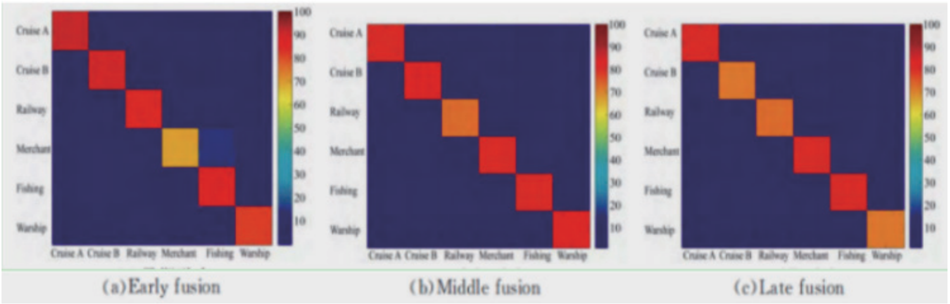

Figure 2Recognition rate matrix of three fusion methods

conclusion

Based on the characteristics of deep convolutional neural network, this paper designs a network model suitable for three band images, and realizes the efficient fusion of three band images. In the aspect of feature extraction and selection, firstly, based on alexnet model, the feature vectors of convolution layer and connected layer are selected, and then the importance of feature vectors is sorted by using ventilator method to reduce the dimension of feature vectors, so as to improve the accuracy of data. In the selection of fusion mode, early fusion, intermediate fusion and late fusion are used. The characteristics of different stages are different, so the fusion recognition rate is also different. In practice, the multi band image database is constructed by using the actual images, and several existing image recognition methods are compared. Experiments show that no matter which fusion method, the final recognition rate is higher than that of a single frequency band.

References:

[1] Ru Fei, litieyingReview of artificial neural network system identification [j]Software guide, 2011,10 (03): 134-135

[2] Wei Na, Ben Kerong, Zhang Linke, pangyunfuResearch on the analysis of underwater vehicle source contribution based on neural network system identification [c]/ / theoretical computer science professional committee of the Chinese computer society. Proceedings of the 2005 National Academic Conference on theoretical computer scienceJournal of computer science, 2005:207-209

[3] YanghuaijiangPerformance analysis of neural network system identification and adaptive control for photodynamic systems [j]Optical precision engineering, 1998 (02): 37-44

[4] Ran Qiquan, Li Shilun, duzhiminOptimized reservoir geological model established by neural network system identification theory [j]China offshore oil and gasGeology, 1996 (01): 33-37

[5] Liuxinghua, liuxianying, Hu ZeResearch on system identification based on RBF neural network [j]Modern electronic technology, 2004 (24): 57-59

[6] LijianxinDesign and Simulation of PID control based on neural network system identification [j]Electronic technology and software engineering, 2021 (08): 122-125

[7] Huang Haocai, Huang Yijian, Yang GuanluNeural network system identification based on LM algorithm [j]Modular machine tools and automatic processing technology, 2003 (02): 8-10+13

[8] Yuhaibo, Ma CuihongSystem identification and application based on diagonal recurrent neural network [j]Microcomputer information, 2007 (31): 216-

217+172
[9] LiuhaifengResearch on system identification method based on neural network [d]Xi'an University of Electronic Science and technology: 2007
[10] Xuxianfeng, Zhang Li, Lang bin, Xia ZhenResearch on face recognition algorithm based on improved twin convolutional neural network with perceptual model [j]Acta electronica Sinica, 2020,48 (04): 643-647
[11] Li Nan, caijianyong, Li Ke, Cheng Yu, zhangmingweiConvolutional neural network face recognition algorithm based on multi perception structure [j] Computer system applications, 2020,29 (02): 157-162
[12] Wangweimin, Tang Yang, Zhang Jian, Zhang YiqiuFace recognition algorithm based on convolutional neural network feature fusion [j]Computer and digital engineering, 2020,48 (01): 88-92+105
[13] Yu YichunFace recognition based on neural network algorithm [j]Electronic technology and software engineering, 2019 (24): 247-248
[14] ShenjiyunAnalysis of face recognition technology based on neural network deep learning algorithm [j]Information technology and informatization, 2020 (05): 228-230
[15] Zhang HanResearch on the application of BP neural network algorithm in face recognition [j]Software, 2020,41 (05): 193-197