Combination of Neural Network and Linear Filtration for Objects Detection

Adilbek K. Shakenov

Institute of Automation and Electrometry, Novosibirsk, Russia, adil.shakenov@ngs.ru

Abstract Several approaches to the use of neural networks for the detection of objects on spatially inhomogeneous backgrounds are considered. Implemented a method for constructing a classifier for detecting objects directly from the observed fragments. An approach is proposed, which consists in a combination of the method of optimal linear filtering and convolutional neural networks. It is shown that the applied approach allows reducing the probability of a false alarm while maintaining the probability of detecting an object.

Keywords: object detection and recognition, convolutional neural networks, machine learning, small-sized objects

1 Introduction

The problem of detecting small-sized, low-contrast objects has been actively studied for the past decades [1]. Research in this area remains relevant, as evidenced by the large number of works on this topic published in subsequent years. When small-sized, low-contrast objects are detected, their shape and size correspond to the hardware function of the system and do not contain enough information for reliable detection. An important feature of the detection algorithms for such cases is the need to evaluate and exclude the underlying background from consideration. The most effective approach for this is the spatio-temporal filtering of image sequences [2]. However, in some cases, due to the peculiarities of the geometry of the survey or the computational limitations of the data processing system, it is necessary to evaluate and suppress the background from one image. There is a known approach to solving the problem under consideration, which makes it possible to obtain an optimal linear filter for the case of a stationary background with a known covariance matrix [3]. Various algorithms for estimating and filtering the background according to the observed local neighborhood are actively developed and applied, for example, bilateral [4], median [5] filtration, optimal linear prediction [6,7], and other heuristic methods [8-11]. The optimal linear filtering method was developed under the assumption that the statistical properties of the background are the same throughout the frame field. This assumption may not hold for a wide range of observed backgrounds. This determines the relevance of the search for new approaches to solving the problem of detecting small-sized, low-contrast objects on spatially inhomogeneous backgrounds.

Recently, methods of recognition and detection of learning neural networks have been actively developed [12]. Examples of use in small objects can be found in [13-16]. In neural networks, a fairly large number of intermediate features are used in the process of processing fragments, so it can be expected that their use will improve the results of linear filtering precisely on spatially heterogeneous backgrounds. In addition, the ability to train the network directly from the observed data makes it possible to easily adapt this approach to a large number of background and observed objects

2 Problem Statement

It is necessary to develop an algorithm for detecting objects on heterogeneous backgrounds, which improves detection characteristics compared to the optimal linear filtering algorithm using trained neural networks.

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-2534/37_short_paper.pdf

3 Detection of objects with training in observable fragments

One of the ways to use neural networks to detect objects is to train a classifier that characterizes each fragment of the observed image as containing an object, or only a background. The size of the processed fragment is chosen equal to the size of the image of the object. The detection procedure with this approach consists in sequentially picking through all fragments of the image and checking them for the presence of an object using a trained classifier. For detection, we used a three-layer convolutional neural network, schematically depicted in Figure 1.

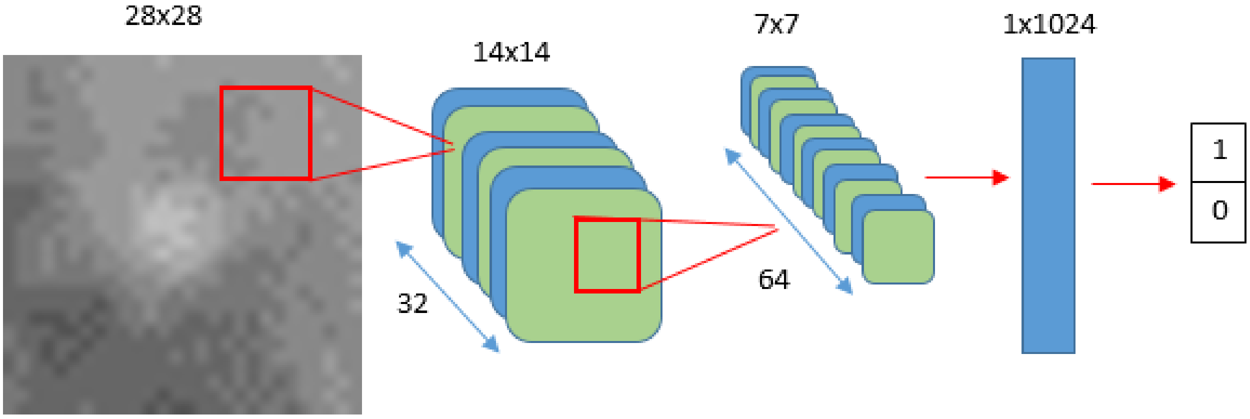

Figure 1. Neural network

The first network layer convolves with 32 different filters of size 9x9 and reduces the size of the resulting images by half. The reduction is carried out by selecting the largest element from a neighborhood of 2x2 pixels. The second network layer similarly performed convolution with two filters of size 9x9 values and halving the size of the output arrays. The third layer converts the resulting data array into one feature vector containing 1024 elements. The resulting feature vector is then characterized as containing or not containing an object.

4 The combination of optimal linear filtering and neural network

Detecting objects using the method described above is rather computationally difficult, since each image fragment must be processed by a neural network, which contains a cascade of a significant number of filters. The optimal linear filtering method gives good enough results for a wide range of real backgrounds, while if the covariance matrix of the background is estimated in advance, the calculation consists in filtering with a single linear filter. Thus, the idea arises at the first stage of processing to use the optimal linear filter, and then apply the trained neural network. The registered image can be represented in vector form as follows:

$$f_{src} = f_o + n,$$

where f_o is the vector of the object, n is the vector of correlated noise (background). If K is the noise covariance matrix, then the linear filter m, optimal in the sense of increasing the signal-to-noise ratio, has the form [3]:

$$m = K^{-1} f_o.$$

In practice, the matrix K, as a rule, is not known. In this work, we used a numerical estimate of the matrix K obtained directly from the input images of the background. Having thus calculated the linear filter, further processing can be carried out according to the scheme shown in Fig. 2.

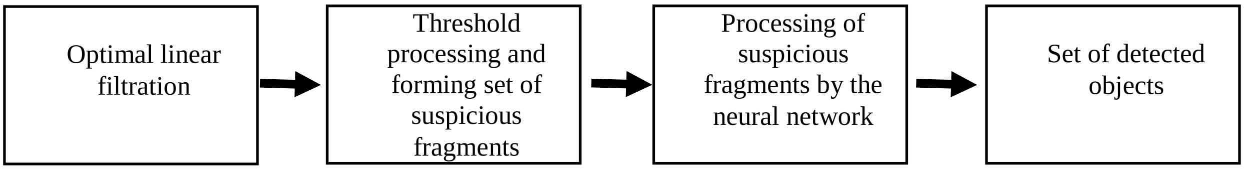

Figure 2. Combination of linear filtration and neural network

5 Source data and network training

For the experiments, we used images of the Earth from the Electro L-1 satellite available in the public domain on the Internet [17]. In the work, point objects are considered, the dimensions and image shape of which are determined by the system's hardware function. The shape of the object was modeled using the Gauss function, the additive method of applying the object was applied. To train the network to recognize fragments after the filtering procedure, the data was obtained as follows. A significant number of objects were applied to the original image at a distance several times the size of the objects. Image fragments containing objects were saved and used to train the network. A filter was built and the image was filtered with applied objects, as well as the original image without an object. From the processed image, a brightness threshold was selected that determines the probability of detecting an object and false alarm. Fragments were selected on the original image containing no objects, the response on the filter in which exceeded the threshold value (false detected fragments). These fragments were subsequently used in training the neural network as examples of a background containing no objects.

6 Experimental results

To compare the effectiveness of the optimal linear filtering and neural network, the following experiment was carried out. One randomly selected background image was used to train the classifier; further comparison of the algorithms was carried out on other background images. As an object, we used a Gaussian function with a maximum intensity equal to one standard deviation of the background and parameter σ equal to 3. To obtain a test set of fragments, 14,000 objects were applied to the background images. Fragments with printed objects were used to assess the likelihood of detecting objects. To assess the likelihood of false alarm, 14,000 fragments of the same size were cut out from the original image in arbitrary places. The resulting sets were processed by a trained classifier. The processing results of several background images are shown in table 1.

Table 1. Comparison of neural network and optimal linear filtering

Texture	**Detection probability**	**False alarm**	
		Net	OLF
Background 1	0,9993	0,0057	0,0011
Background 2	0,9998	0,0061	0,0007
Background 3	0,9996	0,0085	0,0004
Background 4	0,9997	0,0055	0,0005

The “Net” column of Table 1 shows the false alarm values when using the detection algorithm based on the trained classifier, and the “OLF” column - when using the optimal linear filtering. The table shows that in the experiments the neural network showed results worse than the optimal linear filtering. Most likely this is due to the fact that in the absence of a clear methodology, it is quite difficult to choose a training set to obtain a result close to optimal. The data presented show that both algorithms give stable results when processing various input images.

Table 2 shows the results of an algorithm that combines optimal linear filtering and a neural network. In the previous experiment, all fragments of the processed image were fed to the input of the neural network; in this experiment, only fragments suspicious of the presence of an object according to the results of linear filtering. In this case, the “Background 4” image was used with 900 printed objects. Experiments with other backgrounds gave similar results. At the first stage, the optimal linear filtering algorithm was used, and the threshold values were selected that give the detection probabilities indicated in the table in the columnα_1. Each threshold value defines two sets: A_1 - the set of correctly detected objects, and B_1 - the set of false detected fragments that do not contain objects. Denote N_{obj} - the number of all applied objects, N_{pix} - the number of all pixels in the image. Using the sets A_1 and B_1, the probabilities of detection and false alarm were estimated using the optimal linear filtering algorithm shown in the column “OLF1”. The probability of detection was estimated asα_1 = | A_1 | / N_{obj}, the probability of false alarm β_1 = | B_1 | / N_{pix}. Then the sets A_1 and B_1 were fed to the input of the neural network. The result is two sets: A_2 - the set of objects correctly classified by the neural network and B_2 - the set of false fragments of the incorrectly classified neural network as objects. The efficiency of the neural network when processing sets A_1 and B_1, is shown in the column "NeuralNet". It indicates the values α_N = $|A_2|$ / $|A_1|$ and β_N = | B_2 | / B_1, |. The total detection probabilities and false alarms for the detection method considered are given in the “OLF + Network” column. It indicates the values α_C = | A_2 | / N_{obj} and β_C= | B_2 | / N_{pix}. To compare the proposed approach, false alarm probabilities were measured using the optimal linear filtering algorithm for the detection probabilities indicated in the column α_C. These values are given in the "OLF" column and are designated as β_0.

Table 2. Combination of neural network and optimal linear filtering

OLF 1		NeuralNet		**OLF+Net**		**OLF**
α_1	β_1	α_N	β_N	α_C	β_C	β_0
0,999	1,17*10^{-3}	0,968	0,097	**0,967**	**1,11*10^{-4}**	**1,59*10^{-4}**
0,98	1,93*10^{-4}	0,969	0,404	**0,95**	**7,56*10^{-5}**	**1,27*10^{-4}**
0,96	1,38*10^{-4}	0,971	0,480	**0,932**	**6,62*10^{-5}**	**1,07*10^{-4}**
0,94	1,16*10^{-4}	0,976	0,5	**0,918**	**5,61*10^{-5}**	**9,92*10^{-5}**
0,92	1,0*10^{-4}	0,979	0,545	**0,901**	**5,29*10^{-5}**	**9,11*10^{-5}**
0,90	9,03*10^{-5}	0,980	0,564	**0,882**	**4,94*10^{-5}**	**8,13*10^{-5}**

When comparing the values of β_C andβ_0, it can be seen that the proposed approach allowed us to reduce the probability of false alarm by 40-60 percent, with the same probability of detection. The nature of the changes in the values of $\alpha_N$and β_N shows that the results of detection using a neural network correlate with the results of detection by the optimal linear filtering algorithm. To obtain lower values of α_1, it is necessary to use a higher threshold at the stage of threshold processing, which gives sets A_1 and B_1, containing fragments with a higher intensity of the response to the filter. Since the main sign of the presence of an object is an additional registered intensity, it can be assumed that the set of true objects with a higher intensity A_1 becomes easier for correct recognition, and the set of false fragments with a higher intensity B_1becomes more complicated. This can explain the nature of the changes in the quantities $\alpha_N$and β_N. The decrease in $\beta_C$with respect to β_0 is most likely due to the fact that the neural network uses additional features to the filter response that can be used to improve the final results.

7 Conclusion

In the experiments performed, the direct use of a neural network to classify fragments in the considered range of detection probabilities did not improve the results obtained by the optimal linear filtering method. At the same time, the ability to effectively use a combination of optimal linear filtering and a neural network has been shown. Because of applying the proposed approach, the detection efficiency of objects was increased; the probability of false alarm was reduced by 40-60 percent with the same probability of detecting an object. Further research may be aimed at more precise adjustment of network parameters and the use of large amounts of data in the learning process.

References

[1] Gong Cheng, Junwei Han A survey on object detection in optical remote sensing images // ISPRS Journal of Photogrammetry and Remote Sensing. 2016. 117. P. 11-28

[2] Kirichuk V.S., Kosykh I.P., Popov S.A., Sinelschikov V.V. Suppression of a quasistationary background in a sequence of images by means of interframe processing // Avtometriya, 2014, vol. 50, No. 2. P. 3 - 13. Suppression of a quasistationary background in a sequence of images by means of interframe processing // Avtometriya, 2014, vol. 50, No. 2. P. 3 - 13.

[3] Pratt W. K. Digital image processing: PIKS Scientific Inside. PixelSoft, Inc. Los Altos, California p. 662

[4] Tae-Wuk Bae, Kyu-Ik Sohng Small Target Detection Using Bilateral Filter Based on Edge Component // Springer, J Infrared Milli Terahz Waves 2010 Vol. 31, p. 735–743.

[5] Deshpande S.D., Er M.H., Ronda V., Chan Ph. Max-mean and max-median filters for detection of small-targets // Proc. SPIE 3809 (1999): p. 74 – 83

[6] Soni T., Zeidler R., Ku W. H., “Performance evaluation of 2D adaptive prediction filters for detection of small object in image data” IEEE Transactions on Image Processing 1993 2 (3), p. 327–340.

[7] Ffrench P. A., Zeidler J. R., Ku W. H. Enhanced detectability of small objects in correlated clutter using an improved 2-D adaptive lattice algorithm // IEEE Transactions on Image Processing 1997

[8] Hong P., Wang C., Zhang Z. Weak point target detection in the complicated infrared background // Proc. SPIE 2011, Vol. 8200, International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology,

[9] DONG Yu-xing, LI Yan, ZHANG Hai-bo Research on Infrared Dim-point Target detection and Tracking under Sea-Sky-Line Complex Background // International Symposium on Photoelectronic Detection and Imaging 2011: Advances in Infrared Imaging and Applications. Proc. of SPIE Vol. 8193 (2011)

[10] Ivanov V.A., Kirichuk V.S., Kosykh V.P., Sinelshchikov V.V. Features of detection of point objects in images formed by a matrix receiver // Avtometriya, 2016, t.52, No. 2, p. 10-19

[11] Шакенов А. К. Алгоритмы подавления фона в задаче обнаружения точечных объектов по изображениям // Автометрия, 2014, т.50, №4, с. 81 – 87.

[12] Weibo Liua, Zidong Wanga, Xiaohui Liua, Nianyin Zengb, Yurong Liucd, Fuad E.Alsaadi A survey of deep neural network architectures and their applications // Neurocomputing. 2017. 234. P. 11-26

[13] Shangnan Zhao, Yong Song, Yufei Zhao, Yun Li, Xu Li, Yurong Jiang, Lin Li Infrared dim small target segmentation method based on ALI-PCNN model // Proc. SPIE. 2017. 10459. P 104590A-1– 104590A-9

[14] Junhwan Ryu, Sungho Kim Small infrared target detection by data-driven proposal and deep learning-based classification // Proc. SPIE. 2018. 10624. P. 106241J

[15] Zunlin Fan, Duyan Bi, Lei Xiong, Shiping Ma, Linyuan He, Wenshan Ding Dim infrared image enhancement based on convolutional neural network // Neurocomputing. 2018. 272. P. 396–404

[16] Peng Zhang Jianxun Li Neural-network-based single-frame detection of dim spot target in infrared images // Optical Engineering. 2007. 46. P. 076401

[17] Electro-L / Earth from space. http://electro.ntsomz.ru/ (accessed February 1, 2019)