

Development and use of feature models of anthropogenic objects in thematic processing of space imagery

Olga V. Grigoreva¹, Denis V. Zhukov¹, Evgeny V. Kharzhevsky¹ and Andrey V. Markov¹

¹Military Space Academy of A.F. Mozhaisky, Saint-Petersburg, Russia

Abstract

The article describes the problem of automated recognition of the anthropogenic elements of landscape. The recognition is based on the aerospace data in the optical range of the spectrum and a feature model of an object, consisting of the geometric and reflectance characteristics. Using this model, we formed training samples for a convolutional neural network. There is a real example of the practical implementation of the model in identification of the aviation objects.

Keywords

Convolutional neural networks, feature model of object, model adaptation, chain code.

1. Introduction

At present, convolutional neural networks (CNN) are widely used in the automated decoding of space imagery. These technologies help confidently detect and recognize both natural and anthropogenic objects of landscape [1, 2]. However, this approach has its drawbacks. The main one is that neural network technologies usually require a sufficiently large volume of a training sample, which is not always possible to form within experimental data. The designated problem can be solved if the images generated according to the so-called feature models of the object will be applied as a training sample.


2. Feature models

At A.F. Mozhaisky's academy we developed and successfully tested a CNN training method to detect anthropogenic objects, which uses the model data including two basic reference components:

- the geometric characteristics of an anthropogenic object;
- the reflectance characteristics of an anthropogenic object and related backgrounds.


Creating models of anthropogenic objects, we used a contour of figure in the form of a four-connected chain code as a shape reference. For this purpose, adjacent elements of the

SDM-2021: All-Russian conference, August 24–27, 2021, Novosibirsk, Russia

 vka@mil.ru (O. V. Grigoreva)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

object boundaries are connected by vectors in the vertical and horizontal directions [3]. The direction of each vector is encoded with a number. The vector module determines the scale (level of detail) of contour. As a result, the formalized description of the contour is a set of direction codes obtained by tracking the object boundary clockwise.

The reflectance references of anthropogenic objects and related backgrounds consist of two groups of features:

- the color histograms for each component (red, green, blue), which indicate the spatial distribution of pixel brightness over the image surface;
- the spectral signatures of the object and the background in the form of mathematical expectations and standard deviations (SD) of the spectral reflectance coefficient (SRC) and their derivatives (indices, signs of the spectral signature shape, etc.) [4].

In the developed approach, the reflectance features are used for image segmentation into the object and the background at the very first stage of processing. On the other hand, the first group of features is valid for panchromatic and color photographs, but is not always informative for image segmentation. For some objects, the only way to isolate the areas of their potential location is to process multi- or even hyperspectral data [5].

After segmentation the object is detected on a segmented binary image by a trained CNN. Additionally, the segmented image undergoes some morphological operations: opening (or closing) of polygons, filling of one-pixel holes and elimination of the object shadows influencing each other (when the object has elements of different heights).

The required sample volume for CNN training is achieved through the generation of images using feature models adapted to different monitoring conditions.

3. Adaptation of the feature model to the monitoring conditions

The reflectance references are recalculated into pixel brightness values for each color component. The pixel brightness depends on the magnitude of the output of the CCD element. Its value, expressed in the number of photoelectrons accumulated during the exposure time t_n of shooting, is calculated with the following formula:

$$H_{el} = \frac{E(\lambda) \cdot t_n \cdot S \cdot \lambda}{h_{PL} \cdot c},$$

where S is the area of matrix element, m^2 ; $h_{PL} = 6.26 \cdot 10^{-34}$ J·s is Planck's constant; $c = 299792458$ m/s is the speed of light in vacuum; $E = 0.25K^{-2}e_{sum}(\lambda)P\tau_o\tau_s\mu[(r + r_d) + k_p(r + r_d)]$ is the image irradiance, W/m^2 ; K is the lens f-number; τ_o , τ_s is the transmission coefficients of the optical system and light filter; μ is the receiver sensitivity; e is the total spectral irradiance at the equipment entrance; r and r_d are the spectral reflectance coefficients of the object and the background; k_p is the light scattering coefficient of radiation in the monitoring equipment; P is the spectral transmittance of the atmosphere.

Geometric references should be adapted to different values of the linear resolution of the image. At the same time, the linear resolution depends not only on the technical properties of

the camera and the ballistic parameters of the spacecraft orbit, but also on the illumination of the scene.

In general, the linear resolution on the ground is determined with the formula:

$$L = \frac{M}{2R_c} = \frac{H_f}{2f \cdot R_c},$$

where $M = H_f/f$ is the image scale; R_c is the camera resolution with the contrast of the object/background pair (C), reaching the entrance pupil of the camera lens; H_f is the monitoring height, m; f is the focal length of the camera, mm.

The contrast of the object and the background, entering the entrance pupil of the lens, is given by the expression:

$$C = \frac{E_o - E_\phi}{E_o + E_\phi},$$

where E_o and E_ϕ are the integrated illumination of the object and the background in the focal plane (at the entrance to the optical system) in a given wavelength interval, W/m^2 .

After calculating the linear resolution, it is necessary to adapt the geometric reference of the object to a specified scale. The traditional way to solve this problem is to reduce a highly detailed image with rasterization. At this point, the original and the scaled elements of the image are represented by the areal objects. The color of the scaled element is calculated as a proportion of the colors of original element, the area of which intersects the area of the scaled element. The consequence of this approach for segmented images is the appearance of a fuzzy boundary of the adapted segment and the issue of probability of a pixel belonging to an object, which is unacceptable in the geometric analysis.

In the vector description of the object geometry, the scale is set by the modulus of vectors that make up the contour of figure. However, in relation to the objective, if the modulus of vectors of the chain code is proportionally decreased, it will not result in loss of detail in the original image. Therefore, we developed a special conversion algorithm for the chain code to adapt the geometric reference to the linear resolution of the image.

The chain code is adapted by means of vector algebra using vector addition rules. The input data for the algorithm are the original chained code and the scaling parameter, which defines the necessary increase of the modulus of vector of chained code. The parameter is calculated as a ratio of the image resolution to the original scale of the reference data. Object drawings or its highly detailed images should be used as the initial data for a geometry reference of the object. In this case, there is no need to increase the scale of the contour (increase its detail).

In the algorithm, all vectors of the original chain code are considered as unit vectors (the modules of vectors are taken to be equal to 1), but have a different direction. Then the objective is to obtain a contour consisting of vectors, whose modulus equals a given scaling parameter. For this, unit vectors are sequentially added until the modulus of the resultant equals to or exceeds the scaling parameter. The group of vectors is replaced by one vector with a discrete direction (a vector can have only a horizontal or vertical direction). The direction of the new (resulting) vector is determined by a minimum difference between this vector and the result of the unit vectors addition. The calculated difference is added to the sum of the next group of unit vectors. The described operation is repeated until the new longer vectors replaced all the

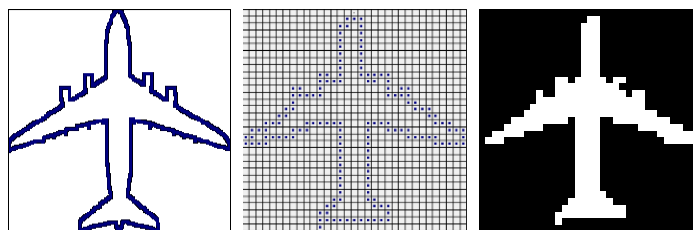


Figure 1: The result of the adaptation of the chain code to the monitoring conditions.

unit vectors. An example of the adaptation of the chain code to the monitoring conditions is shown in Figure 1.

4. Conclusions and example

The data set collected with the help of adaptation can provide a sufficient sample for CNN training in the detection of anthropogenic objects. A high efficiency of the described method was experimentally confirmed by the success in the detection of aircrafts at an airfield. YOLO 3

Example of the original image



Example of the result of segmentation and recognition



Figure 2: The result of using a feature model of an aircraft for its recognition.

was used as a CNN. In this case, as related backgrounds we considered typical airfield pavements: asphalt and concrete. The geometric references of the aircraft (chain codes) were calculated upon their digital drawings. The spectral reflectance coefficient of the objects and related backgrounds were extracted from a special database developed at the Academy. It's based on the results of the Academy's own field experiments and in-flight spectrometric measurements carried out for over thirty years. Figure 2 shows an example of the feature model of an aircraft used for its recognition.

Therefore, the obtained data indicate that thematic processing of images with neural networks gives a high reliability results.

Additionally, there is an important advantage of the proposed approach: it allows recognizing the desired object even without its image references, and the described list of references of anthropogenic objects and related backgrounds supports the necessary level of quality for CNN training.

References

- [1] Jamali A., Mahdianpari M., Brisco B., Granger J., Mohammadimanesh F., Salehi B. Comparing solo versus ensemble convolutional neural networks for wetland classification using multi-spectral satellite imagery // *Remote Sensing*. 2021. Vol. 13. P. 2046. DOI:10.3390/rs13112046.
- [2] Ivanov E.S., Tishchenko I.P., Vinogradov A.N. Multispectral image segmentation using convolutional neural network // *Current Problems in Remote Sensing of the Earth from Space*. 2019. Vol. 16. No. 1. P. 25–34.
- [3] Gonzalez R.C., Woods R.E. *Digital image processing*. 2th edition. 2002. ISBN-10:0201180758.
- [4] Grigorieva O.V., Markov A.V., Ivanez M.O., Zhukov D.V. Methods of preparation of formal etalon features for target detection using hyperspectral remote sensing data // *Vth All-Russian Scientific and Technical Conference with International Participation "Focal Problems of Space-Rocket Hardware" (Vth Kozlov Readings)*. Samara: JSC SRC Progress, 2017. Vol. 1. P. 281–286.
- [5] Grigoryeva O.V. Observation of forest degradation using hyperspectral data aerial and satellite sensing // *Earth Observation and Remote Sensing*. 2014. No. 1. P. 43–48.