№2 (5) 2024

INTERNATIONAL SCIENCE REVIEWS

AIU

Natural Sciences and Technologies series

ISSN: 2707-4862

ASTANA INTERNATIONAL UNIVERSITY

INTERNATIONAL SCIENCE REVIEWS
Natural Sciences and Technologies series

Has been published since 2020

№2 (5) 2024

Astana

Editorial address: 8, Kabanbay Batyr avenue, of.316, Nur-Sultan,
Kazakhstan, 010000
Tel.: (7172) 24-18-52 (ext. 316)
E-mail: natural-sciences@aiu.kz

International Sciense Reviews NST - 76153
International Science Reviews
Natural Sciences and Technologies series
Owner: Astana International University
Periodicity: quarterly
Circulation: 500 copies

CONTENT

GRNTI 28.23.15

RECONSTRUCTION OF GEOMETRIC MODELS OF OBJECTS FROM SATELLITE IMAGES BASED ON ARTIFICIAL NEURAL NETWORKS

Shalbai T., Kaldarova M., Nazyrova A., Sultangaziyeva A., Kussepova L.

Astana International University, Astana, Kazakhstan.
E-mail:kmiraj8206@gmail.com

Abstract: Segmentation of natural objects, such as soil and vegetation types, is possible using methods that operate on the spectral brightness of individual pixels, determined by phytocenological, humification and other mechanisms of their homeostasis. Segmentation of objects of artificial nature, such as houses, railways and highways, requires knowledge about a large number of spectral parameters, the structure of the object and its surroundings. Modern approaches to extracting three-dimensional information about man-made objects can be conditionally divided into traditional (based on classical processing methods) and neural network (data mining). The capabilities of the first methods have exhausted themselves under the conditions of the increasingly complex structure of the earth's surface and the need for automation of data processing. The second group of methods is being actively developed at the present time, its possibilities are unlimited and their quality is directly related to the volume of the training sample. The article discusses a method for constructing three-dimensional models of earth surface objects, based on a two-stage (integral and local analysis) and multiscale approach and assuming that the initial data is pre-processed by the method of increasing the resolution, which is characterized by a representative decrease in the resolution of the training sample images.

Keywords: object segmentation, data mining, integral and local analysis, multiscale approach, U-Net

INTRODUCTION

One of the most important tasks of satellite image processing is the development of algorithms for automatic selection and segmentation of objects of natural nature, which is possible with the use of relatively simple classical methods of digital processing.

In recent years, the de facto standard for image processing tasks is the use of convolutional artificial neural networks, which allow taking into account the context and invariably occupy the first places among image segmentation and classification algorithms.

METHODS AND METHODOLOGY

Authors [1-3] a method for constructing three-dimensional models of rigid objects on the Earth's surface from a single satellite image on the example of railway infrastructure objects is proposed. The method consists of step-by-step processing of satellite images with the sequential use of two convolutional neural networks.

Due to the growth of computing power and the creation of large image databases, effective training of deep verification neural networks has become possible. Convolutional neural networks have achieved high success in the task of processing objects. So, every year, since 2012, the Image Net Large Scale Visual Classification Challenge (ILSVCC) competitions are won by teams using the SNC. In scientific research, convolutional neural networks are widely used for classifying and detecting objects in images [12, 13]. A trained convolutional neural network creates a class label and generates a bounding box of the area in which the object class is located in the image.

Scientists from all over the world are trying to solve the problem of vegetation classification using all available information from remote sensing of the Earth and applying various methods and algorithms to it [14-16]. However, it should be noted that only a few works were devoted to the detection of damaged trees using satellite images or images of unmanned aerial vehicles. At the same time, these works present the results of detecting only a few categories of tree damage (for example, healthy and damaged trees) using classical image classification methods.

At the moment, different methods of preparing a control sample for training a convolutional neural network are used, for example: manual preparation of the Image Net dataset, which has been carried out since 2010 to the present, involving more than fifty institutions [5], a method of preparing a new data set using convolutional neural network methods [6], building a test sample using the Image Data Generator function of the Keras module written in Python. Also, to build a set of image data, the functions of the Chainer module, developed by Preferred Networks in partnership with IBM, Intel, Microsoft and Nvidia, an image geoprocessing tool using the ArcGIS program, etc. are used. All this requires a lot of time to build and label data, the cost of computing power, money, if you prepare a test and control sample using ready-made software products.

Deli et al. [7] developed a new SNA-based method for classifying four classes of land cover (crops, house, soil and forest) on remote sensing images. They submitted 100 images for each class to the SNA for its training.

M. Längkvist, A. Kiselev, M. Alirezaie, A. Loutfi [11] proposed the classification and segmentation of satellite information, distinguishing five classes: vegetation, land, road, parking, railway, construction and water supply, using convolutional neural networks on images of a small city for a complete, fast and accurate pixel-by-pixel classification. The authors selected the parameters and analyzed their impact on the architecture of the neural network model. They also found better performance of the convolutional neural network model compared to object-oriented classification methods (Object-Based Image Analysis, OBIA), while the maximum classification accuracy using convolutional neural networks reached 94.49%.

Some researchers have used neural network classifiers to recognize plant species and life cycle in images obtained from a digital camera and a mobile phone camera. Thus, the authors of the work [8] created a new convolutional neural network capable of distinguishing 22 plant species in color images with an accuracy of 86.2%. To do this, the authors used data sets that differ in lighting, image resolution and soil type. In the article [9], a new architecture of a convolutional neural network was proposed for the classification of plant types growing at agricultural plants. To evaluate the effectiveness of this approach, the results of the created convolutional neural network model were compared with the Support Vector Machine (SVM) method. The classification accuracy of the developed convolutional neural network was significantly higher than the results of the classical method, reaching 97.47%.

A number of papers present the results of the use of convolutional neural networks in the analysis of satellite images. In particular, several studies have carried out the classification of vegetation on satellite images using convolutional neural networks [10].

MATERIALS AND METHODS

Here are the main advantages of a convolutional neural network relative to other machine learning methods:

* In general, the convolutional neural network in many tasks gives significantly higher accuracy of recognition and classification of objects in images compared to other methods.

• Compared to a fully connected neural network, a convolutional neural network has a much smaller number of trained weights, since one core of weights is used entirely for the entire image, instead of creating its own personal weight coefficients for each pixel of the input image. This pushes the neural network during training to generalize the input information, and not to memorize each image presented to it by a huge number of weight coefficients, as a fully connected neural network does.

* Convenient parallelization of calculations, and, consequently, the possibility of implementing algorithms for working and training the network on GPUs;

* Relative stability to rotation and shift of the recognized image;

* Training using the classical method of error back propagation.

The main disadvantage of the SNA is the presence of a significant number of parameters configurable by the researcher. At the same time, it is often unclear which optimal parameter values should be chosen for a specific task, taking into account the limitations on available computing power. The variable network parameters include: the number of layers, the dimension of the convolution core for each of the layers, the number of cores for each of the layers, the step of the core shift during layer processing, the need and number of subdiscretization layers, the degree of dimension reduction by them, the dimension reduction function (selection of maximum, average, and others), the transfer function of neurons (activation function), the presence and parameters of a fully connected neural network at the convolutional output. All these parameters significantly affect the result, but are chosen by researchers empirically. A number of verified and efficient network configurations have been developed, but in general there are no recommendations for building a network for a new task.

Within the framework of this method, a U-Net-based topology was applied, expanded and supplemented by the authors for the tasks of interpreting aerospace images. Due to the need to solve several problems within the same model, to reduce the training time and ensure the consistency of the results, instead of an ensemble of independently trained models used traditionally, the authors introduced the following modifications of the topology of the neural network model and the process of its training: the encoder block is left unchanged; for each of the subtasks, separate decoder blocks and loss functions are implemented; In order to optimize RAM consumption, a separate gradient descent operation is performed for each of the decoders during model training [1-3].

At the first stage of processing (integral analysis), a satellite image is segmented using a neural network to select a set of objects of specified classes. At the second stage of processing (local analysis), a local analysis of the image areas identified by the results of the first stage of processing is performed using a neural network. The results of the second stage of processing are used to evaluate the parameters of the three-dimensional model of the object. The possibilities of the method are shown by the example of processing a satellite image of the railway infrastructure. The following classes of informative areas are considered: a building, a wall edge, a roof edge, a building shadow, railway infrastructure, a car, a highway; rails, pillars and shadows from pillars (taken as reference objects for evaluating scaling coefficients in various directions). An example of the application of the developed method for identifying typical railway infrastructure objects and for subsequent evaluation of the parameters of a three-dimensional model of a building is given.

Integral analysis-identification of the main properties and relationships of objects in the image:

- Objects belong to thematic classes: buildings, structures, fences, railway tracks, etc.;
- Pixels belong to spatial classes: roofs and walls of buildings, the surface of the earth;
- Whether pixels belong to the shadows of objects;
- Pixels belonging to a significant neighborhood of objects;
- Grouping image pixels into objects and extracting relevant information (meaningful context).

Local analysis is the processing of each identified object in order to build its three-dimensional model. It is based on the use of the results of the first stage, depending on which the object is subjected to various processing procedures:

- unique objects (determined by pre-known spatial positions) are not processed, if available, ready – made 3D models are used;
- typical objects (determined by the thematic class of the object determined at the stage of integral analysis) are subjected to a specialized (for this type of object) procedure for determining properties;
- for the remaining objects, the geometry restoration procedure is performed.

The schematic diagram of the topology of a neural network model based on UNet with several decoders looks like this [4].

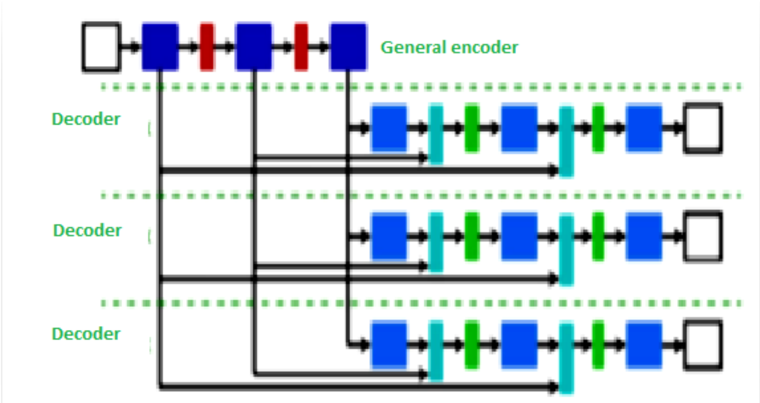

Figure 1. - Schematic diagram of the topology of a neural network model based on G-Net with several decoders [4]

Modifications of the U-Net topology were applied:

- the encoder block was left unchanged;
- separate decoder blocks and loss functions are implemented for each of the subtasks;

- in order to optimize RAM consumption, a separate gradient descent operation is performed for each of the decoders during model training.

General requirements for training and test satellite images:

1. Seasonality of shooting (lack of snow cover)
2. Proper visibility conditions – shooting in the daytime, minimizing shooting artifacts (false pixels, interlacing, etc.)
3. Correctness of satellite survey data (it is advisable and timely to investigate updated objects and processes on the earth's surface)
4. Constant and high spatial resolution within the satellite image (higher resolution for railway automation, telecommunications or telemechanics objects than for railway platforms and station buildings)
5. Geo-linking images to railway infrastructure objects
6. Geo-referencing of images (linking images to a single coordinate projection – the universal transverse Mercator UTM projection)

The variability of the image characteristics is allowed, which consists in:

- The possibility of using images with different sets of spectral channels
- The presence or absence of metadata for individual images
- Full or partial visibility of objects in images
- The possibility of using a small-volume training sample.

The choice of informative classes is as follows:

1. Building (polygon, c1)
- Roof (polygon, c2)
- Roof edge (polyline, c3)
- Wall edge (polyline, c4)
- Shadow of the building (polygon, c5)
2. Railway infrastructure (landfill, c6)
- Railway track (polygon, c7)
- Pole (polyline, c8)
- Shadow of the pole (polyline, c9)
- Wagon (polygon, c10)
3. Highway (polygon, c11)
4. Other classes:
- Walls (polygon, c'=c1-c2)
- Excluded objects (polygon, c'')
- Background (c''')

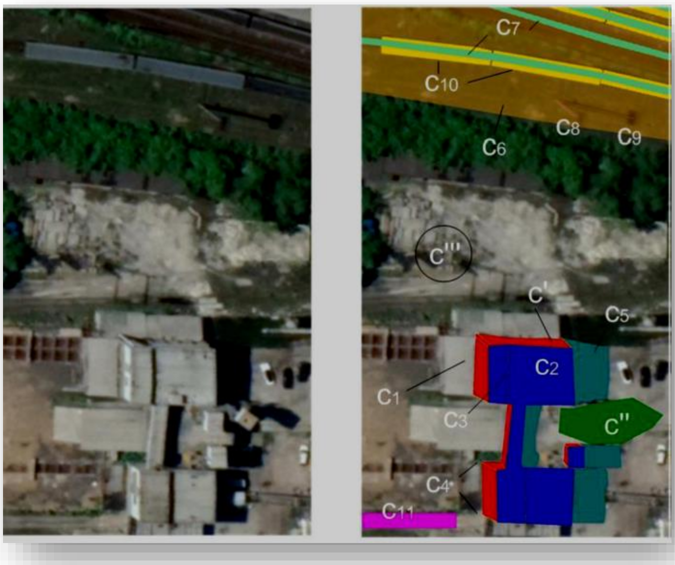

Figure 2. - Image and markup, respectively

Dimensions of typical objects:

R3=1.52 m-railway track width, class c7

R4=11.35 m-the height of the post, class c8

We note a number of features of these classes that make their markup for training quite a complex and time-consuming task:

1. Self-similarity of areas
2. Adding regions
3. Obscuring areas
4. Distortion of areas
5. Intersection of areas
6. Poor visibility

Rules for building markup:

1. The assumption rule
2. The separability rule
3. Exclusion rule

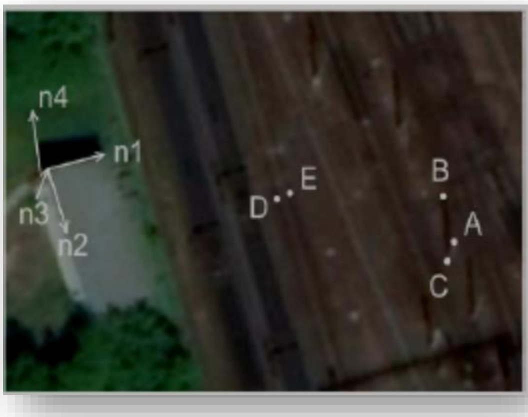

Figure 3. - Building a geometric model of an object based on raster areas [4]

CONCLUSION

Thus, we agree with the multiscale method proposed by the authors [4] for constructing three-dimensional models of objects of the Earth's surface based on a separate orthotransformed satellite image, consisting of two stages: integral and local (object-by-object) analysis, each of which includes a composition of traditional and neural network methods.

Integral analysis is aimed at localization of objects and relevant information: thematic and spatial classes of individual pixels, shadows, and so on. Local analysis - to extract the data necessary to determine the geometry of the object and their interpretation.Neural network models are built on the basis of the UNET full-convolutional topology, expanded and supplemented by the authors for the tasks of interpreting aerospace images. Local analysis supports the construction of models of both typical objects with low variability according to pre-known features, and buildings or structures of rectangular spatial configuration.

In the article, we described an approach to reconstructing a three-dimensional model of rigid objects from a single image without using metadata based on four informative classes identified by the results of machine learning.

REFERENCES

1. Richter A. A., Murynin A. B., Shlupikov V. A. Features of integrating remote sensing data for restoring three-dimensional models of rigid objects from satellite images / / Materials of the International Scientific and Practical Conference "Achievements of Science in 2019", December 27, 2019, the Center for Scientific Development "Big Book" (Moscow). URL: http://conf.rse.geosmis.ru/ DOI 10.21046/18DZZconf-2020a

2. Gvozdev O. G., Murynin A. B., Richter A. A. Complex of applied solutions for the construction and training of artificial neural networks for semantic segmentation of aerospace images of an arbitrary channel-spectral structure in conditions of a shortage of training data / / Materials of the 19th All-Russian Conference with international scientific participation: Mathematical methods of pattern recognition (MMRO-2019, Moscow), Russian Academy of Sciences, 2019. - pp. 344-348. URL: http://conf.rse.geosmis.ru/ DOI 10.21046/18DZZconf-2020a

3. Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation / Computer Vision and Pattern Recognition (cs.CV) 18 May 2015. URL: https://arxiv.org/abs/1505.04597.

4. Richter A. A., Gvozdev O. G., Murynin A. B., Kozub V. A., Kosheleva N. V. Restoration of geometric models of railway infrastructure objects from satellite images based on artificial neural networks / / Materials of the 18th All-Russian Open Conference "Modern problems of remote sensing of the Earth from space". Moscow: ICI RAS, 2020. p. 41. URL: http://conf.rse.geosmis.ru/files/books/2020/8118.htm. DOI 10.21046/18DZZconf-2020a

5. Russakovsky, O., Deng, J., Su, H. et al. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis 115, 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y

6. Guirado, E.; Tabik, S.; Alcaraz-Segura, D.; Cabello, J.; Herrera, F. Deep-learning Versus OBIA for Scattered Shrub Detection with Google Earth Imagery: Ziziphus lotus as Case Study. Remote Sens. 2017, 9, 1220. https://doi.org/10.3390/rs9121220

7. Deli, Z. Farmland scene classification based on convolutional neural network / Z. Deli, C. Bingqi, Y. Yunong // In Proceedings of the 2016 International Conference on Cyberworlds (CW). – 2016. – P. 159–162. DOI:10.1109/cw.2016.33

8. Plant species classification using deep convolutional neural network. / Dyrmann, Mads; Karstoft, Henrik; Midtiby, Henrik Skov.In: Biosystems Engineering, Vol. 151, No. November, 11.2016, p. 72-80. https://doi.org/10.1016/j.biosystemseng.2016.08.024

9. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition. *Proceeding of the IEEE*. 1998. p. 2278-2324. Google Scholar

10. Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. Remote Sens. 2017, 9, 22. https://doi.org/10.3390/rs9010022

11. Längkvist, M. Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks / M. Längkvist, A. Kiselev, M. Alirezaie, A. Loutfi // Remote Sensing. – 2016. – № 8. – P. 329 URL: file:///C:/Users/%D0%9F%D0%9A/Downloads/remotesensing-08-00329.pdf (дата обращения 14.07.2021). doi:10.3390/rs8040329

12. Bass L.P., Kuzmina M.G., Nikolaeva O.V. Svertochnyie neironnyie seti s glubokim obucheniem v zadachah obrabotki giperspektralnih sputnikovyh dannyh//Басс Л.П., Кузьмина М.Г., Николаева О.В. Сверточные нейронные сети с глубоким обучением в задачах обработки гиперспектральных спутниковых данных // Препринты ИПМ им. М.В.Келдыша. 2018. № 282. 32 с. doi:10.20948/prepr-2018-282 URL: http://library.keldysh.ru/preprint.asp?id=2018-282

13. Воронцов, К.В. Машинное обучение [Электронный ресурс] / К.В. Воронцов // 2019. – Режим доступа: http://www.machinelearning.ru (дата обращения: 11.08.2021).

14. Safonova A.N., Dmitriev E.V. Classification of agricultural crops from middle-resolution satellite images using gaussian processes based method, J. Sib. Fed. Univ. Eng. technol., 2018, 11(8), 909-921. DOI: 10.17516/1999-494X-0113

15. Dmitriev E.V. Combining classifiers in the problem of thematic processing of hyperspectral aerospace images / E.V. Dmitriev, V.V. Kozoderov, A.O. Dementyev, A.N. Safonova // Optoelectronics, Instrumentation and Data Processing. – 2018. Volume 54, Issue 3. – P. 213–221. https://ui.adsabs.harvard.edu/link_gateway/2018OIDP...54..213D/doi:10.3103/S8756699018030019

16. Khovratovich T.S. Forest change detection based on sub-pixel estimation of crown cover density using bitemporal satellite data / T.S. Khovratovich, S.A. Bartalev // Space Research Institute RAS. Sovr. Probl. DZZ Kosm. – 2019. – № 16. – P. 102–110. URL: http://jr.rse.cosmos.ru/article.aspx?id=2037&lang=eng. DOI: 10.21046/2070-7401-2019-16-4-102-110