Accuracy evaluation of automated object recognition using multispectral aerial images and neural network

Conference Paper · August 2018

DOI: 10.1117/12.2502905

3 authors:

Volodymyr Hnatushenko
Dnipro University of Technology

Dmitry Mozgovoy
Dnepropetrovsk National University

Volodymyr V. Vasyliev
EOS DATA ANALYTICS, Inc.

Accuracy Evaluation of Automated Object Recognition Using Multispectral Aerial Images and Neural Network

Dmitriy Mozgovoy*[a,b], Volodymyr Hnatushenko[a,b], Volodymyr Vasyliev[a,b]
[a]EOS Data Analytics, 1906 El Camino Real, Menlo Park, CA 94027, USA;
[b]Oles Gonchar Dnipro National University, 72 Gagarina av., Dnipro, Ukraine

ABSTRACT

Automated classification of urban land cover is a classical problem in remote sensing. A methodology of accuracy evaluation of automated object recognition using of sub-meter spatial resolution multispectral aerial images and neural network is proposed. The methodology is applied for 5 land cover classes detection from visible and infrared images using a multilevel convolution neural network (CNN). Image processing results are analyzed. It is shown that the recognized object boundaries are delineated with sufficiently high accuracy and classes are well separated.

Keywords: multispectral aerial images, convolution neural network, accuracy evaluation, map updating.

1. INTRODUCTION

1.1. Relevance

At present, for regular maps of cities, multispectral aerial images are increasingly used. EOS Data Analytics company developed a web-service for automated recognition of building, vegetation, water bodies, etc. in cities using aerial images of sub-meter spatial resolution in visible and infrared ranges of electromagnetic spectrum in order to increase the efficiency and reliability of updating maps of cities [1]. With the development of high-resolution remote sensing images, the requirements of classification are getting higher and higher. The classification of aerial scenes gives birth to a series of studies on the deep learning in the field of remote sensing. The popularity of convolution neural network in this field is primarily because it is highly adaptive, no need of priori knowledge, and can be computed parallelly [2-6]. The accuracy and reliability of CNNs depend on the network's training and the selection of operational parameters [7].

1.2. Existing problems

Accuracy evaluation of the automated object recognition by multispectral aerial images is the final stage of testing of the developed system before putting it into operation. For comparison of the obtained result with other existing ones it is necessary to determine the following:
- How to evaluate correspondence;
- With what data to compare the obtained result;
- How to evaluate the quality of the obtained result (the obtained evaluation of difference includes not only the error proper of the tested result, but also it is stipulated by the error of the test data).

1.3. Objectives

The objective of the work is accuracy evaluation of the automated object recognition using multispectral aerial images and a neural network.

The validation of a web service of the automated classification of multispectral aerial images lies in the defining the basic indicators of classification quality:
- classification accuracy (the percentage of unrecognized and falsely recognized classes);
- reproducibility (the repeatability of results in various test cases);
- sustainability (no significant deviations in the results of the object recognition should appear in the input data or during the setting of a processing procedure).

* dmitriy.mozgovoy@eosda.com; phone 3 8097-3621887

2. AUTOMATED OBJECT RECOGNITION

2.1. Input data

Multispectral aerial images of high spatial resolution registered in 2012, USA (California, Contra Costa), are used as the input data for 7 testing areas (example is shown on the Figure 1). The main characteristics of the aerial images are as follows:
- spatial resolution – 15…30 cm;
- radiometric resolution – 8…11 bits;
- number of spectral channels – 4;
- coordinate System - UTM / WGS-84;
- data format – GeoTIFF.

Figure 1. The fragment of aerial image for one testing area (RGB composite of visible bands)

2.2. Image processing technology

The main stages of multispectral aerial images processing are:

- preliminary processing of data for the reference area (selection of the area of interest, mosaicking of images and cropping by kmz-file);
- spectral synthesis of visible bands RGB composites for reference region and formation of vector layers of the training samples for each class in interactive mode;
- preliminary processing of data for the control area (selection of the area of interest, mosaicking of images and cropping by kmz-file);
- classification of visible and infrared images using a multilevel convolution neural network (CNN) and generated training samples in automated mode;
- morphological filtering, object classification and vectorization of each class in interactive mode (example is shown on the Figure 2);
- spectral synthesis of RGB composites for the reference area and the formation of vector layers of the reference classes in interactive mode (example is shown on the Figure 3);
- calculation of the objects characteristics of each class, recognized in the control area in automated mode and their comparison with the reference classes obtained in interactive mode;

- calculation and visualization of errors and accuracy estimation of the automated recognition (example is shown on the Figure 4).

Figure 2. The of vector layers of the recognized classes.

Figure 3. The of vector layers of the reference classes.

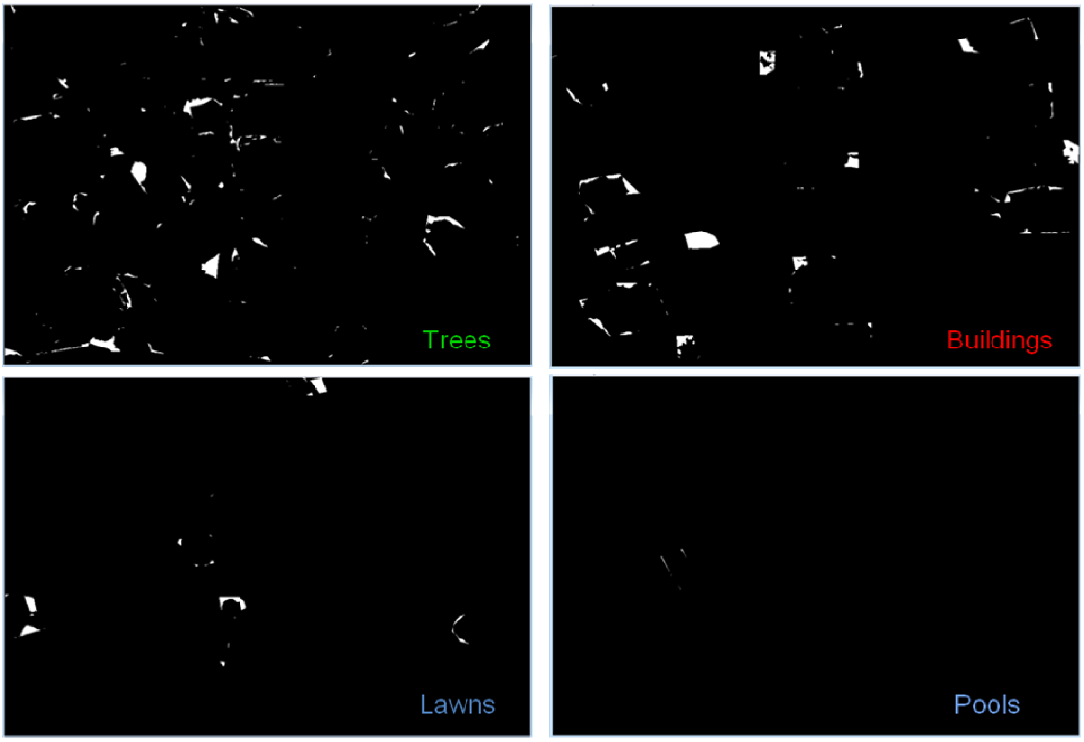

Figure 4. The visualization of recognition error for 4 classes.

2.3. Recognizable object classes

The neural network was taught to recognize 5 classes [4]:
- class 1 - high vegetation (deciduous trees and shrubs, coniferous trees).
- class 2 - low vegetation (fields, lawns).
- class 3 - buildings and constructions (single- and multi-storey).
- class 4 - water objects (ponds, pools, etc.).
- class 5 - shadows of tall objects (buildings, trees) – an axillary class that is used to exclude the shadowed areas and aimed to reduce the classification errors, which may appear due to inaccurate recognition of the objects shadowed by trees or buildings.

As soon as the model data had been classified only by one common class of vegetation the evaluation of classification accuracy was prepared only for classes 1 and 2 that were united into one class of vegetation.

3. EVALUATION OF RECOGNITION ACCURACY

3.1. Model accuracy evaluation

The results of terrestrial measurements of vegetation, considered as the model, were appeared quite erroneous due to the following reasons:
- a time gap between taking images and making terrestrial measurements;
- the complexity of accurate recognition of a projective surface of trees;
- georeferencing system error for all model data layer (shift arising because of using of custom datum);
- instrumental and methodical errors of measurements.

To reduce these errors, manual correction of the vector vegetation layer (the model data) using satellite images of the specific area has been accomplished. The evaluation of model accuracy has been accomplished by an expert method. The average accuracy of the model is:
- 3…5% without manual correction of the model data
- 1…2% with manual correction of the model data.

3.2. Methodology of accuracy evaluation

The most widely used methods of validation of Earth remote sensing classification outcomes are the following:

- comparison of results with the results of synchronous surface observations and measurements carried out immediately at the time of imaging;

- comparison with the results of automated classification by certified software products for the same purpose;

- comparison with the outcomes of manual classification carried out by operators and evaluated by an expert group (this method is used for comparatively small volumes of data or for a limited set of test areas, which are to be distributed over the territory of the research as evenly as possible).

Taking into account the above, the accuracy of classification evaluation is accomplished by comparison of automated classification with manual corrected surface observations results.

3.3. Applied metrics

The metrics that have been used for the quantifying accuracy of the automated classification are the following:

- confusion matrix (a number of unrecognized class pixels, a number of falsely recognized class pixel, total result accuracy);
- statistics (Kappa coefficient, a regression or standard error);
- compliancy matrix for several classes (the accuracy can be low due to the transition of class boundaries);
- compliancy matrix with fuzzy boundaries of the class; in this case an algorithm of class boundaries designation may vary.

In this case the following well-known indicators of accuracy classification have been chosen: confusion matrix and Kappa coefficient.

For the accuracy evaluation of one class the matrix of classification errors has been used (Figure 5).

		Condition Positive	Condition Negative	
EOSDA Analytics Outcome	Decision Outcome Positive	True Positive (TP) Hit	False Positive (FP) False Alarm	Positive Predictive Value = TP / (TP + FP) Precision
	Decision Outcome Negative	False Negative (FN) Miss	True Negative (TN) Correct Rejection	Negative Predictive Value = TN / (FN + TN)
		Sensitivity = TP / (TP + FN) Hit Rate	Specificity = TP / (FP + TN) True Negative Rate	

Figure 5. The matrix of classification errors for 1 class.

The overall evaluation of the accuracy of automated classification is determined by the formula:

$$\text{Overall accuracy} = \frac{TP + TN}{N}. \quad (1)$$

A confusion matrix for several classes is an instrument that applies a cross-tabulation for the evaluation of correlation among the values of matching classes obtained from different sources. The sources are the checked bitmaps (for the automated classification) and a more accurate, supporting data source (for the manual classification).

The class names of the checked data set are indicated on one matrix axe, the model classes used for checking are indicated on another matrix axe. The cases of consistency of calculated classes and actual data are represented on the main diagonal of the matrix (an accurate classification).

The example of confusion matrix for 5 classes (A, B, C, D, E) is shown in Table 1.

Table 1. Confusion Matrix

		Manual classes					
		A	B	C	D	E	∑
Neural classes	A	n_{AA}	n_{AB}	n_{AC}	n_{AD}	n_{AE}	$n_{A\_}$
	B	n_{BA}	n_{BB}	n_{BC}	n_{BD}	n_{BE}	$n_{B\_}$
	C	n_{CA}	n_{CB}	n_{CC}	n_{CD}	n_{CE}	$n_{C\_}$
	D	n_{DA}	n_{DB}	n_{DC}	N_{DD}	n_{DE}	$n_{A\_}$
	E	n_{EA}	n_{EB}	n_{EC}	n_{ED}	n_{EE}	$n_{E\_}$
	∑	$n_{\_A}$	$n_{\_B}$	$n_{\_C}$	$n_{\_D}$	$n_{\_E}$	N

The overall accuracy, kappa coefficient, confusion matrix, errors of commission, errors of omission, producer accuracy, and user accuracy are calculated. Confusion matrix using either a ground truth image or using ground truth ROIs and both produce an output similar to the following example.

The overall accuracy is calculated by summing the number of pixels classified correctly and dividing by the total number of pixels. The ground truth image or ground truth ROIs define the true class of the pixels. The correctly classified pixels are on the main diagonal of the confusion matrix table, which lists the number of pixels that were classified into the correct ground truth class. The total number of pixels is the sum of all the pixels in all the ground truth classes.

3.4. Accuracy indices

To evaluate the accuracy of a particular calculated class, the number of correctly classified pixels of this class should be divided on the total number of pixels of the class according to verification data. This index is called producer's accuracy as it displays how well the classification result of this class coincides with the verification data. For class A producer's accuracy is:

$$A_p = \frac{n_{AA}}{n_A}. \quad (2)$$

A similar index can be calculated for the model class if the number of correctly classified pixels is divided on overall pixel number of this class according to the verification data. This index is called user's accuracy as it shows the user of the classification how likely is that this class coincides with the classification results. For class A user's accuracy is:

$$A_u = \frac{n_{AA}}{n_{A\_}}. \quad (3)$$

A sum of values of the diagonal elements shows a total number of correctly classified pixels. The ratio of this number to a total pixel number on the N matrix is called an overall classification accuracy and it is usually expressed in percent:

$$\text{Overall accuracy} = \frac{n_{AA} + n_{BB} + n_{CC} + n_{DD} + n_{EE}}{N}. \quad (4)$$

The off-diagonal elements show the discrepancy of calculated and actual classes (classification errors).

The kappa (κ) coefficient measures the agreement between classification and ground truth pixels. A kappa value of 1 represents perfect agreement while a value of 0 represents no agreement.

$$\kappa = \frac{N\sum_{i=1}^{n} m_{i,i} - \sum_{i=1}^{n}(G_i C_{i,i})}{N^2 - \sum_{i=1}^{n}(G_i C_{i,i})}, \quad (5)$$

Where :

i is the class number;

N is the total number of classified pixels that are being compared to ground truth;

$m_{i,i}$ is the number of pixels belonging to the ground truth class i, that have also been classified with a class i (i.e., values found along the diagonal of the confusion matrix);

C_i is the total number of classified pixels belonging to class i;

G_i is the total number of ground truth pixels belonging to class i.

4. MAIN RESULTS

4.1. Testing outcomes

Analysis of the results of image processing showed a sufficiently high accuracy of the delineation boundaries of recognized objects and a good separation of classes.

A comparative assessment of the accuracy of automated classification of buildings, vegetation, and water objects using multispectral aerial images and neural network for three testing districts has been accomplished.

The indices of classification accuracy of testing district №1 (for 5 classes):
- kappa index = 0,83;
- overall accuracy = 90,39%.

For other testing districts the indices of classification accuracy (for 4 and 5 classes) were within the following limits:
- Kappa index = 0,81…0,95;
- overall accuracy = 90,02…92,45%.

The above basic indicators confirmed the high classification quality (classification accuracy, reproducibility and sustainability).

4.2. Directions for further research

Currently, work is under way to modernize of methodology of automated object recognition using of sub-meter spatial resolution multispectral aerial images. Main directions for further research include:

- small mobile objects recognition (car parsing or road traffic monitoring);
- complex processing of lidar and multispectral aerial images;
- automated change detection for different objects or classes.

5. CONCLUSION

The above-mentioned indices of accuracy of buildings, vegetation, and water objects automated classification using multispectral aerial images serve as a proof of the following:

- the used methodology is rather effective;
- the methodology can be used for solving applied tasks.

The biggest advantages of the automated land cover classification compared to manual terrestrial measurements are:

- automated classification is much faster than manual measurements (calculated mainly by the timing of aerial photography);
- it can be applied to any area coverage including closed and hard to get areas;
- no expenses for measure operators, transport etc.;
- lower price of classification, especially on big territories (determined basically by the cost of aerial imagery);
- more accurate results, absence of man factor (intentional data misrepresentation);
- high level of automation, the ability to use in pair with existed GIS resources;
- the ability of automated creation, filling an updating spatial databases.

The developed methodology provides a significant increase of the efficiency and reliability of updating maps of cities.

Due to the high degree of automation, the proposed methodology can be implemented in the form of a geo-information web service, functioning in the interests of a wide range of public services and commercial institutions.

REFERENCES

[1] https://eos.com. (19 January 2018).

[2] Linlin Cao. "Application of convolutional neural networks in classification of high resolution remote sensing imagery". Science of Surveying and Mapping, 2016(9), pp. 170-175.

[3] Castelluccio M., Poggi G., Sansone C. "Land use classification in remote sensing images by convolutional neural networks". Journal of Molecular Structure Theochem, 2015, 537(1), pp. 163-172.

[4] Hnatushenko V.V., Mozgovoy D.K., Serikov I.Ju., Vasyliev V.V. "Automatic vegetation classification using multispectral aerial images and neural network". System technologies. Regional Interuniversity collection of scientific works. – Edition 6 (107). - Dnepropetrovsk, 2016. - P.66-72.

[5] Hnatushenko V.V., Hordiiuk D.M. "Neural Network and Local Laplace Filter Methods Applied to Very High Resolution Remote Sensing Imagery in Urban Damage Detection". 2017 IEEE International Young Scientists Forum on Applied Physics and Engineering YSF-2017. October 17-20, 2017, Lviv, Ukraine. P.363-366. doi:10.1109/YSF.2017.8126648

[6] Hu F., Xia G.-S., Hu J., and Zhang L. "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery". Remote Sensing, vol. 7, no. 11, pp. 14680–14707, Nov. 2015.

[7] Radovic M.; Adarkwa O.; Wang Q. "Object Recognition in Aerial Images Using Convolutional Neural Networks". J. Imaging, 2017, 3, 21. doi:10.3390/jimaging3020021

AUTHORS' BACKGROUND

Your Name	Title*	Research Field	Personal website
Dmitriy Mozgovoy	assistant professor	digital image processing, machine learning and remote sensing applications	
Volodymyr Hnatushenko	Dr. Sc. (Dr.-Ing. habil.), full professor	pattern recognition, SAR image processing, machine learning and remote sensing applications	http://www.dnu.dp.ua/en/department_of_automated_systems_of_information_processing
Volodymyr Vasyliev	Phd candidate	image processing, machine learning and remote sensing applications	https://eos.com/company/#team