Implementation convolution neural network model method for solve real time objects detection task with geospatial approach use

A.A. Makhmutov
LLC
Innovations technology for solution
in geographical information systems
Ufa, Russia
e-mail: arthur.makhmutov@gmail.com

Abstract

In this paper, considering the implementation of the model of a convolutional neural network for solving real-time object detection problems using the approach of geoinformation technologies for processing and mapping of detected objects as geospatial units. Also this article covers issues related to the peculiarities of training, software and hardware implementation of the model of the convolutional neural network. The relevance of this article is to identify the behavior of detectable objects in different conditions on the investigated spatial sections using the convolutional network as deep learning method, therefore, it allows identifying predicative models, the use of which is necessary for solving the problems for urbanized entities infrastructure control.

Keywords: Real Time Nearest Remote Sensing, Deep Learning, Convolutional Neural Network, GIS, Machine Learning, Recognition, Object Detection

1. Introduction

At the present time, the directions of deep learning and geoinformation technologies are dynamically developing as independent industries in the information technology world. The use of methods of machine learning in geoinformation technologies is widely used for the processing of remote sensing data of the Earth. There are more than ten machine learning methods that are successfully implemented in leading geographic information systems, such as ENVI, GRASS, QGIS, ArcGIS, Erdas Imagine [1]. The main methods of machine learning used to work with remote sensing data include decision tree methods, support vector machines, back propagation neural networks, and percussive perceptrons [2].

Proceedings of the 6th All-Russian Scientific Conference "Information Technologies for Intelligent Decision Making Support", May 28-31, Ufa - Stavropol, Russia, 2018

As practice shows, to obtain more accurate results of processing space images, neural networks of backward propagation are used [3]. Conventional neural networks are used to recognize objects, including remote sensing data of high and ultra-high spatial resolution. The use of the method of convolutional networks has found wide application in various industrial directions, including autonomous systems. The advantage of this method is the rapid processing of this and more accurate recognition of objects. In this paper, we present the results of a study of the use of a convolutional neural network to solve the problem of identifying objects in real time on the sites under investigation using the QGIS software platform. Recognition of objects in real time becomes popular in 2018 and begins to be widely used to solve a wide range of tasks, including capture of criminals. In this connection, the term "Real time nearest remote sensing" is introduced in this article.
Real time nearest remote sensing is a method of remote sensing of objects and phenomena in the investigated area, using deep learning methods to collect information about objects and phenomena, as well as predicting their behavior in real time.

2. Related researches

As know, the dawn of convolutional neural networks began from the development of Alex Krizhevski in 2009 [4]. From this time began to develop direct for visual recognition of objects at images and video files [5]. The research in the direct of prediction of actions of objects became actual [6]. Collecting information about behavioral models of objects allowed to make a forecast about subsequent actions or movements [7]. Models that are used to predict actions are widely used in computer vision, compression of video files. Also, these models are used to recognize voice commands and train voice assistants, which are used in modern consumer electronics, automotive systems, smartphones. Currently, there are several models of convolutional networks that are of value to the scientific and industrial communities, including AlexNet, DarkNet, Fast R-CNN PSPNet [8].

3. Method

3.1. Equipment

For making the calculations, used hardware platform with a graphics card built on the NVIDIA GFORCE GTX 1050TI architecture with 768 CUDA cores.

3.2. Model choose

For solving the research problem, the Darknet- based model was chosen for several reasons. At first, this model is well optimized for recognizing a large number of objects. At second, this model has high accuracy of recognition and prediction. At third, this model has a high recognition speed using already existing data sets, as well as individually created ones.

3.3. Dataset

In the process of the research tested two datatested, Yolo9000 and Coco, were originally used to recognize the image objects received from the outdoor surveillance cameras with 720p resolution, the best recognition results were shown using the Coco dataset (Fig. 1. Yolo9000 and Coco results of classification).

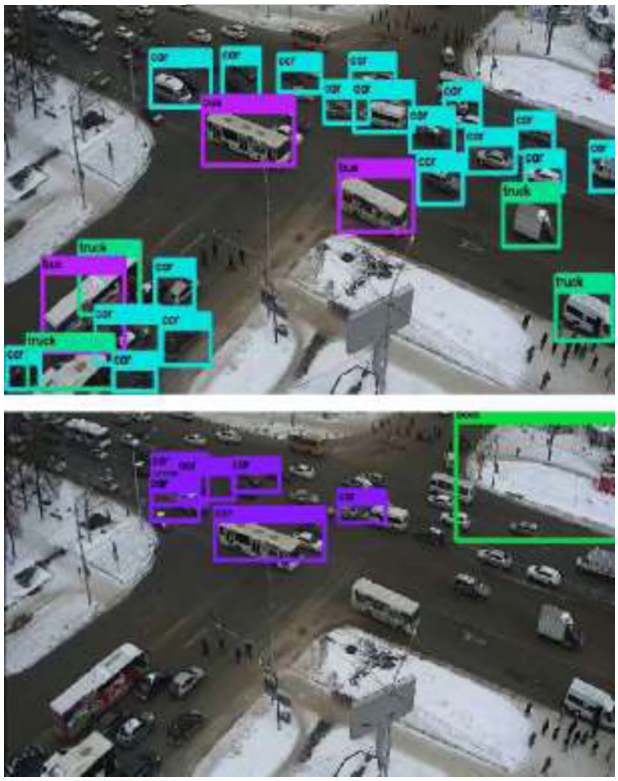

Fig. 1. Yolo9000 and Coco results of classification: a) Classification result for Coco dataset; b) Classification result for Yolo9000 dataset

But to solve the problem of recognition people by sex, age and vehicles by type and brand, necessary to create individual dataset, using combined data. To create individual dataset, used the data obtained from the sensors of external video cameras, and data of object models made with photo cameras used (Fig. 2. Dataset sample for one class).

Fig 2. Dataset sample for one class

The need to create a custom dataset consisted of the climatic conditions of the region where the study was being carried out. in connection with the often falling precipitation, the appearance of the objects of recognition has the property of changing. To create 78 classes, 2000 images were initially used, which were culled and placed in the training sample files. At the next stage, training samples were marked and made annotations for every class.

3.4. Training

The learning process for 78 classes was conducted for 90,000 iterations, the total training time of the neural network was 720 hours, or 30 calendar days. The learning process completed when the average loss values decreased below the 0.100000 value. In the configuration parameters used for the model for the Darknet- based convolutional neural network, the training step was set to 0.001000. After 45000 iterations, the neural network training was monitored using datasets obtained at the output with an interval for each 5000 iterations. Also, to monitor training, the data received from the outdoor surveillance cameras of several sites was used. As the result of the training, the result of recognizing the given classes was achieved with an average accuracy for 98% (Fig. 3. Detecting objects with custom dataset use).

Fig 3. Detecting objects with custom dataset use

4. Programming realization

The software implementation of the Darkned-based convolutional neural network realized in C ++ programming language with use the NVIDIA framework and the CUDA and CuDNN libraries designed for accelerating the computation at the GPU platform. For realization also used computer vision library OpenCV by version 3.4. For implementation in the form of a software solution, the source code was compiled in the C++ dll library that was used in the final compilation. Integration

of the convolutional neural network within the geoinformation system used with the Quantum GIS API (Fig. 4. Programming components, used for task realization).

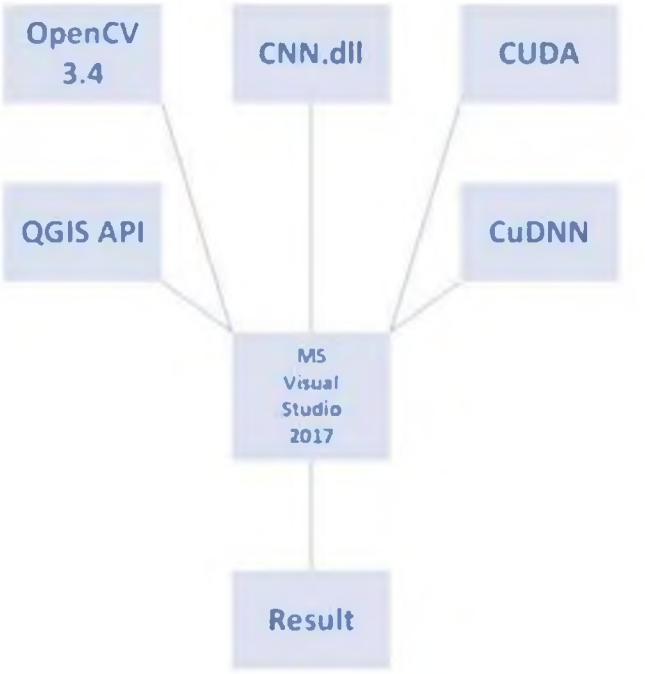

Fig 4. Programming components, used for task realization

5. Results

This research allowed to solve the common implementation problem for use deep learning and geoinformation technologies, this approach contributed to the emergence of a new method of remote sensing. The method of real time nearest remote sensing helps to identify the features of events and phenomena of the study areas for further decision making that will contribute to the emergence or prevention of new phenomena, depending on the specificity of the area under study.

6. Conclusion

- The use of convolutional neural networks solves a large number of tasks related to public safety, transport security, industrial sphere, the identification of objects and phenomena, the prediction of events, the management of unmanned systems.
- The implementation of the method of the real time nearest remote sensing allows to learn about events and phenomena at the moment, this becomes especially important for decision-making.
- The use of the nearest real time remote sensing method in geoinformation systems helps to assess the situation at the study areas, taking into account the real time, geographic location, and the peculiarities of the terrain on which actions are currently taking place.

7. References

1. A. A. Makhmutov. “Use Hamming neural network for remote sensing data classification in GRASS”. 2016
2. Makhmutov A.A., Khristodulo O.I., Gatiyatullin D.I., Musin I.A. Estimation classification accuracy of algorithms based at artificial intelligence methods for remote sensing data of South Ural area Proceedings of the 4-th international conference ITIDS’2016. Ufa State Aviation Technical University 2016; 2: 19-24.
3. O.I. Khristoduloa, A.A. Makhmutova, T.V. Sazonovab. Use algorithm based at Hamming neural network method for natural objects classification. XIIth International Symposium «Intelligent Systems», INTELS’16, 5-7 October 2016, Moscow, Russia. Procedia Computer Science 103 (2017) 388 – 395.
4. A. Krizhevsky. Learning multiple layers of features from tiny images. Master’s thesis, Department of Computer Science, University of Toronto, 2009
5. L. Yao, A. Torabi, K. Cho, N. Ballas, C. J. Pal, H. Larochelle and A. C. Courville. Describing videos by exploiting temporal structure. In ICCV, 2015
6. M. Hoai and F. De la Torre. Max-margin early event detectors. In CVPR, 2012
7. T. Lan, T.-C. Chen, and S. Savarese. A hierarchical representation for future action prediction. In ECCV, 2014
8. Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia. Pyramid Scene Parsing Network, The Chinese University of Hong Kong, SenseTime Group Limited. arXiv:1612.01105v2