Possibilities of deep learning neural networks for satellite image recognition

A N Averkin¹,³ and S A Yarushev²

¹ Scientific Laboratory of Artificial Intelligence, Neuro-technologies and Business Analytics, Plekhanov Russian University of Economics, Moscow, 117997, Russia
² Department of Informatics, Plekhanov Russian University of Economics, Moscow, 117997, Russia
³ Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences

E-mail: averkin2003@inbox.ru

Abstract. The main problem solved in this project is the analysis of big data using a system of computer processing and recognition of satellite images, based on a deep neural network architecture. The goal of the project is to develop methodological, theoretical and practical aspects of building such systems in poorly formalized subject areas, as well as to study the possibilities and advantages of building predictive models for analyzing fresh water reserves and predicting the direction, speed and nature of the spread of large fires using such systems, and assessments of the economic impact of these natural disasters.

1. Introduction

This project is aimed at developing a deep learning system for predicting the distribution of water resources on the planet using satellite data processing. Modern technologies make it possible to receive and update quite often and store processed images in distributed data storages. In this project, it is proposed to build a system that allows analyzing the obtained data in a semi-automatic mode, comparing them with publicly available cartographic information and information received from experts in the field of water resources distribution.

Remote sensing satellite is a modern and effective method of tracking an object or phenomenon in the absence of direct physical contact. There are two methods: active and passive. Probing with the use of passive sensors involves fixing signals emitted or reflected by an object or territory. An example of such a signal can be sunlight, and an example for a method can be digital or film photography, infrared devices. Active sensing implies the emission of signals for scanning. An example would be radar.

Thanks to remote sensing, it is possible to obtain information about hard-to-reach, dangerous and fast-moving objects. For example, remote sensing can improve the chances of preventing and eliminating early fires and outbreaks in hard-to-reach areas where the response time of services is delayed.

Satellite images of natural phenomena are studied by numerous organizations for monitoring and control purposes. In the literature there are a number of works devoted to the analysis and prediction of the development of natural phenomena using neural networks (T.S. Stankevich, Application of convolutional neural networks for solving the problem of operational forecasting of the dynamics of the...
spread of forest fires. The difference between the proposed project and existing storage systems and analysis of satellite imagery data consists in the dynamic capabilities of the proposed system, that is, functionality for computer data generation (photorealistic images of reservoirs with fresh water, forest fires, and many others) for training deep neural networks.

Let us explain this in more detail. Satellite imagery allows you to obtain high-quality images of fresh water reservoirs and forest fires. However, the amount of available satellite images required for full-fledged training of deep neural networks in order to predict the dynamics of fresh water reserves and, especially, the dynamics of the spread of forest fires, is usually insufficient. This is especially true for fires due to their high propagation speed and the fact that the smoke associated with the fire makes it difficult to satellite imagery in the visible spectrum. While it is not difficult to obtain a separate high-quality image of a forest fire, the formation of a series of images suitable for training deep neural networks (covering several days or even weeks of the evolution of the fire) is a serious problem: the satellite is moving in orbit, cloudiness prevents shooting, the air is clouded with smoke. This kind of interference also hinders satellite imagery of freshwater bodies, albeit to a lesser extent.

To overcome this circumstance, within the framework of the proposed project, a system of imitation modeling of the studied objects and phenomena (reservoirs with fresh water, forest fires, landscapes, etc.) will be developed, which will allow generating images missing for training deep neural networks.

A new approach used in this project is the analysis of big data using a computer simulation system for the dynamics of fresh water reserves and the characteristics of large forest fires, based on a deep neural network architecture. The goal of the project is to develop methodological, theoretical and practical issues for the development of such systems in poorly formalized subject areas, as well as to explore the possibilities and advantages of building predictive models for analyzing natural resources using such systems. Various aspects of the decision-making problem based on the results of the work of this predictive model will be considered both with the help of a quantitative assessment of factors within the framework of operations research models, and with the help of a qualitative analysis of the mutual influence of system-forming factors of the studied subject area within the framework of intelligent information systems.

2. The problem

Recently, an increasingly important role is played by satellite remote sensing methods of earth exploration. This is due both to the further improvement of space technology: an increase in the quality, quantity and availability of satellite images, and to the curtailment of aviation and ground-based monitoring methods.

The fields of application of satellite remote sensing are different, they are obtaining information about the state of the environment and, the state of affairs in land use, and forecasting the level of crop yields, and assessing the consequences of natural disasters: earthquakes, floods, forest fires. The use of remote sensing tools is also effective in studying the measure of soil and water pollution, as well as assessing the extent of ice melting in the Arctic and Antarctica or controlling deforestation in the taiga. Among other things, the use of these methods makes it possible to obtain data on the state of the atmosphere, including on a global scale. Due to the fact that the amount of data is huge, the analysis of satellite images can be effective using machine learning methods.

Thanks to this, processing, analysis and compilation of data suitable for subsequent analysis by specialists is carried out. One of the most effective and highly promising methods of obtaining data on the position of various objects on the ground, as well as tracking their changes, is the automatic recognition of satellite or aerial photographs. A major advance in this direction is the move away from manual image segmentation. This problem becomes especially relevant when it is necessary to process large areas of the earth's surface in a short time.

The use of artificial satellites enables scientists to observe current phenomena and track changes, the nature of the development of events. Remote sensing can increase the ability to prevent and eliminate at an early stage fires and ignitions in hard-to-reach areas, where the response time of the Ministry of
Emergency Situations, as well as rescue services, may be delayed. Timely information is one of the main means of minimizing risks and damage caused.

With the development of information technologies and the general increase in the computing power of computers, it became possible to simulate processes of a high degree of complexity related to intelligence, behavior and the mind of a person as a whole, which became a factor in the emergence of a new scientific discipline - artificial intelligence. Artificial intelligence consists of subsections, one of which is machine learning.

Today, artificial neural networks based on deep learning methods cope with the task of recognizing satellite images in the best way. There are a wide variety of types and architectures of neural networks, such as multilayer perceptron (MLP), convolutional neural networks (CNN), recurrent neural networks (RNN), or so-called long short-term memory networks. (long short-term memory - LSTM), each of which is used with varying success to solve specific problems. It is customary to classify neural networks according to several parameters: by the type of input information, the nature of training, and the models of connections between neurons in the network. The most common way to choose a neural network architecture is to study the best practices used by researchers in related tasks. In particular, for solving the problem of recognizing satellite images, the use of convolutional neural networks is best suited, which are the most used tool for solving problems related to image processing.

3. Deep neural networks for image recognition

Deep learning [1] technology, which is seen as a breakthrough in the development of artificial intelligence, delivers the highest accuracy of image recognition and speech, but it is still applicable only to limited types of data. In particular, it has so far been difficult to classify accurately automatically variable time series data coming from the devices connected to the Internet of things.

Machine learning is used not only as a mathematical theory, but also applied in practice, it is an engineering discipline. In most cases, theoretical foundations do not immediately lead to algorithms and methods that show themselves well in practice. To optimize them and improve the quality of work, it is necessary to modify the developed methods to compensate for the deviation of the theory from the practical conditions. Indeed, much of the research in machine learning needs to be tested in practice, running an experiment on simulated or real data that confirms that the method works. Machine learning methods Machine learning as a separate section appeared as a result of the division of the science of neural networks into methods of training networks and types of topologies of their architecture; machine learning also includes methods of mathematical statistics. The following machine learning methods are based on the case of using neural networks, although there are other methods that use the concept of a training sample - for example, discriminant analysis that operates on the generalized variance and covariance of the observed statistics, or Bayesian classifiers. Such classical models of neural networks as the perceptron (multilayer perceptron) are capable of learning in different ways, with or without a teacher, with reinforcement and self-organization. However, many types of neural networks can only be trained one way (like most static methods). For this reason, the classification of machine learning methods may be incorrect, in the case when the sorting is done according to the training method (referring neural networks to a certain type).

Jeffrey Hinton [2] with his graduate students is one of the creators of the "deep learning" method of artificial neural networks. For the first time, they applied their methodology in a competition in the field of pharmaceuticals. The task was to create and apply a program based on a limited set of data on the chemical composition of molecules that could determine the most effective molecules that, in turn, can work as efficiently as possible as a medicine. Another interesting fact is that the team applied for participation in the competition at the very last moment. And the system they developed was created without information about how the object - the molecule interacts with the subject - the target cell, which it must cure.
4. **Capsule neural network**

Capsule neural networks were proposed by Jeffrey Hinton, one of the researchers who proposed to train a neural network using error backpropagation. In 2012, he proposed to train deep neural networks using the backpropagation method, which later allowed for a breakthrough in image recognition problems. In October 2017, he published work [3], in which he presented a new architecture of neural networks called capsule neural networks. This architecture can kind of revolutionize the field of recognition, as they can solve the problems of everyone's favorite convolutional neural networks. The problem with a convolutional neural network (CNN) is that during the training process for image recognition, information about the spatial relationships of all functions is lost. CNN does not remember the position of image elements in space, the same object, but from a different angle it can already be perceived as another image. Therefore, you have to make a huge sample for the same object from all angles. Capsule neural networks can use spatial relationships to solve a similar problem (see figure 1).

![Figure 1: Capsule neural network.](image)

This approach outperformed CNN by a significant margin, reducing errors by 45%. Each of the capsules is designed to detect a specific function of the image so that it can recognize them from different angles. This approach is the newest in the field of artificial neural networks.

5. **Other deep learning applications**

To solve the problem of text classification, a method was proposed [4] based on a modified neural network with learning backpropagation of an error. The method was used to minimize the number of characteristics of the training sample. Mountcastle [5] is one of the leading researchers in the study of the functioning of the human brain. The columns in the brain are connected to each other by lateral connections. Links have positive and negative feedback. In some way they resemble Kohonen's self-organizing maps [6]. Neurons with identical signals are located next to each other, which looks like a self-organizing map.

Implementation of neural network models is possible in many languages programming. When choosing a programming language, the most priority characteristics [7]:

- the computational speed of the language;
- the reliability of the language;
- availability of software libraries or frameworks for the solution machine learning tasks, as well as their level of documentation;
- "Entry threshold" - the difficulty of mastering a programming language.

Today, the leader among programming languages is Python. Assessing this language according to the characteristics described, Python is reliable, is on an approximate level in terms of computational speed with other programming languages, is easy to learn and has a large number of math libraries, data visualization libraries and machine learning frameworks developed in open source. After the release of the language, Python was considered a classic "mathematical" programming language, due to the laconic code and the rich number of ready-made tools. With the development of the field of machine learning, in particular deep learning, Python has gained great popularity due to its mathematical focus.
To solve deep learning problems, there are several large frameworks of different levels, both low and high, such as, for example, TensorFlow. TensorFlow (TF) is an open source library developed and maintained by the IT industry leader Google.

Development in open source allows you to see the full software implementation of all algorithms, and also provides quick interaction (support) of the framework developers with the community of programmers using the issue-tracking subsystem of the GitHub service, on which the source code is posted.

The principle of operation of TensorFlow is to build a graph of operations (or a computation graph) and perform calculations in accordance with the constructed model. The neural network graph includes variables-tensors (variables) and all operations, starting with the input layer, ending with the output layer and the optimizer (learning algorithm).

TensorFlow provides the ability to fully implement the mathematical model described in the previous section, while the library is also suitable for a wide range of machine learning techniques in addition to deep learning. TF has a lightweight TensorFlow-Slim add-on to create an artificial neural network model. Slim is a "wrapper", a compact implementation, for the layers of the neural network, and the initialization of the neuron weights is performed using the algorithms listed in the previous section by the internal capabilities of TF. The flatten function is an operation to go from 3D data obtained on the last pooling (or convolution) layer to 1D data, used to go to the fully connected part of the neural network. Thus, flatten is a flat representation operation, identified in the TensorFlow-Slim add-in as a separate layer type. Along with the implementation of neural network layers, TensorFlow provides an implementation of most of the activation functions. The choice of the activation function for the neurons of each layer is set through an additional parameter of the Slim functions, for example, as in the case of the output layer and its softmax activation function (formula 8). The default activation function for most layers is set to ReLU.

In addition to the implementations of neural network layers, TensorFlow provides the ability to use any of the learning algorithms listed in the previous section. Adam's algorithm is no exception and has its own implementation named AdamOptimizer. AdamOptimizer allows setting all additional parameters: attenuation coefficients B1 and B2, smoothing term E and learning rate coefficient A. In the process of training neural networks, the key definition of TF is sessions (Session object). In the session, data is transmitted to the input and operations are performed, and the session itself stores the current state of each element of the computation graph. Thus, to capture changes such as learning outcomes, you need to save the session.

Session saving can be done at any time, the conditions of saving are set by the programmer independently. Session is saved using the tf.learn.Saver class, and the saved graph model is a separate binary file with the .ckpt extension (checkpoint files). Using a trained neural network model implies loading (restoring) the saved model by the tf.saved_model.loader loader class, searching for the required tensors and performing the required operations. For example, to recognize images, you need to find the tensor the input layer, which is responsible for receiving input data, and the tensor of the output layer, which gives the recognition result, perform the recognition operation and get the result. In addition to the direct use of the trained model, the reconstructed session value can be used to retrain the neural network on new data to obtain results of greater accuracy. In addition to additional training, TensorFlow sessions for convolutional neural networks allow visualization of feature maps of convolutional layers. Since the saved model allows you to select any tensor in the graph, it is not difficult to find tensors responsible for feature maps. Visualization of feature maps of convolutional layers is an effective way of showing where the neural network is "looking" in the image and which key areas allocates.

6. Conclusion
In this paper, we considered machine learning methods for recognizing satellite images. The basics of neural networks, machine learning algorithms and their areas of application were studied. A study of approaches to training neural networks was carried out. The collection and analysis of data, search and study of scientific literature in the subject area was carried out.
7. Acknowledgments

The paper is supported by the scientific-research grant «Intelligent system of satellite data analysis for the forecasting of economic sequence of water resources and fire damages” of Plekhanov Russian University of Economics (Protocol 4, April, the 27th of 2020).

References
[1] Ranzato M A and Szummer M 2008 Semi-supervised learning of compact document representations with deep networks Proceedings of the 25th international conference on Machine learning 792-99
[2] Hinton G E, Osindero S and Yee-Whye 2012 A fast learning algorithm for deep belief nets Neural computation 18(7) 1527-54
[3] Sabour S, Frosst N and Hinton G E 2017 Dynamic routing between capsules Advances in Neural Information Processing Systems 30 (NIPS 2017) Proceedings 3859-69
[4] Yu B, Xu Z and Li C 2008 Latent semantic analysis for text categorization using neural network Knowledge-Based Systems 21 900-904
[5] Mountcastle V 1997 The columnar organization of neocortex Brain 120 701-722
[6] Kohonen T 2001 Self-Organizing Maps (Springer Verlag)
[7] Averkin A N and Yarushev S A 2018 Time series and data analysis based on hybrid models of deep neural networks and Neuro-Fuzzy networks CEUR Workshop Proceedings 2267 453-456