

DEVELOPMENT OF THE PROJECT TO CREATE INFORMATION TECHNOLOGY FOR RECOGNITION TEXT FROM IMAGES USING CNN

Vinnytsia national technical university

Abstract

Most of the information people perceive in the form of images, their automated analysis is a natural and very informative way of describing the environment. As the number of sources of obtaining visual information is rapidly increasing, the task of automated image analysis becomes extremely relevant. The most effective means of automated decision-making in such classification tasks are artificial neural networks, in particular, convolutional neural networks.

Keywords: text, detection, convolutional neural networks

Анотація

Більшість інформації, яку люди сприймають у вигляді зображень, їх автоматизований аналіз є природним і досить інформативним способом опису середовища. Оскільки кількість джерел отримання візуальної інформації швидко зростає, завдання автоматизованого аналізу зображень стає надзвичайно актуальною. Найбільш ефективним засобом автоматизованого прийняття рішень у таких завданнях класифікації є штучні нейронні мережі, зокрема, згортальні нейронні мережі глибокого навчання.

Ключові слова: текст, розпізнання, згортальна нейронна мережа

Introduction

It is proposed to create a project for recognition text messages from images, which is based on the use of modern methods of convolutional neural networks of deep learning. Deep learning is also called hierarchical learning or deep structured learning. The technology received its name because of the fact that the data passes through several layers of the neural network. The results of one layer operation are input data for another. This allows the machines to independently learn and create a hierarchical display of data.

Scientific novelty of the obtained results:

For the first time, a classifier for recognizing the content of a text message through deep machine learning based on the use of technology of convolutional neural networks, which is based on a comparative analysis of the frequency of use of each of the set of criteria for the distribution of text message content, and uses the framework of advanced Caffe machine learning, which increases the accuracy and the quality of text recognition from the image.

Results of research

Consider the typical structure of the convolutional neural network in more detail. The network consists of a large number of layers. After the initial layer (input image with text message), the signal passes through a series of convolutional layers, which contain a convolution layer and a sub-sampler layer. Alternating layers allows you to make a "charts" and from the signs of cards, on each next layer the card decreases in size, but the number of channels increases. In practice, this means the ability to recognize the complex hierarchy features.

With the notion of training associated with a lot of activities, it is difficult to provide a clear definition of this process. Moreover, the learning process depends on the point of view on it. That makes

practically impossible the emergence of any precise definition of this concept. For example, the learning process from the point of view of a psychologist is fundamentally different from learning from the point of view of a school teacher. From the position of the neural network, probably, you can use the following definition:

Training is a process in which the free parameters of the neural network are tuned by simulating the environment into which this network is built. The type of learning is determined by the way of adjusting these parameters.

This definition of the learning process of the neural network involves the following sequence of events:

- The neural network receives stimuli from the external environment.
- As a result of the first paragraph, the free parameters of the neural network are changed.
- After changing the internal structure of the neural network responds to the violation in another way.

There are two conceptual approaches to training neural networks:

- training a neural network with a teacher assumes that for each input vector from the training set there is a necessary value of the original vector, called the target. These vectors form a learning pair. Weights of the network are changed until each input vector receives an acceptable level of deviation of the source vector from the target;

- training a neural network without a teacher is a much more plausible model of learning from the point of view of the biological roots of artificial neural networks. The training set consists of only the input vectors. The algorithm for training the neural network adjusts the weight of the network so that the agreed output vectors are obtained, that is, the presentation of sufficiently close input vectors yielded the same outputs. To teach neural networks without a teacher, the signaling method of training Hebb is used. Figure 1 shown schematic of the neural network training.

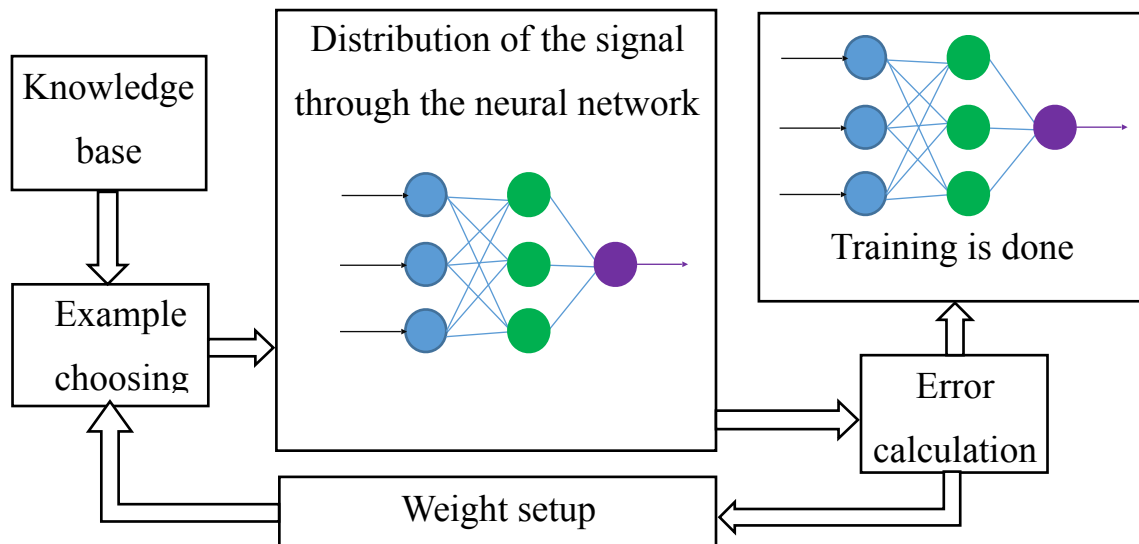


Figure 1 – Schematic of the neural network training.

Mathematically, the learning process can be described in this way. In the process of functioning, the neural network generates an output signal Y , realizing some function $Y = G(x)$. If the network architecture is specified, then the type of function G is determined by the values of the synaptic weights and the displaced network. Let the solution of some problem be a function $Y = F(x)$, given input-output parameters $(X^1, Y^1), (X^2, Y^2), \dots, (X^n, Y^n)$, for which $Y_k = F(X_k)$ with what $(k = 1, 2, \dots, n)$.

The software for the task of recognition was developed in the package for Matlab R2016b applications using the Deep Learning toolbox. A classifier developed on the basis of a convolutional neural

network was trained in images from the California University of Technology. This university is one of the most cited in the used collection of images collected from well-known brands and trends.

It was found that the probability of text recognition from the image is more than 96%. The results of the recognition can be seen in figure 2.

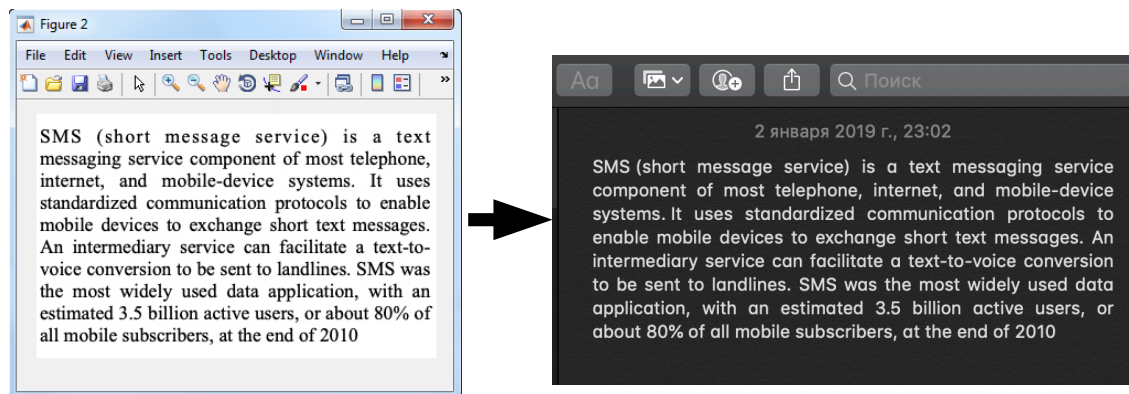


Figure 2 – Recognize text from a downloaded image

Conclusions

The proposed information system should have a high probability of recognizing the text messages in comparison with known software solutions. The developed mathematical-statistical analysis should describe the work of the information system in full.

References

1. Бісікало О.В. Визначення змістовних ознак тексту на основі аналізу зв'язків між лексичними одиницями / О.В. Бісікало, А.І. Лісовенко, О.В. Яхимович, С.С. Траченко // Вісник НТУ «ХПІ». Серія: Механіко-технологічні системи та комплекси. – Х: НТУ «ХПІ», – 2015. – № 21 (1130). – С. 83–89. – Бібліогр.: 10 назв. – ISSN 2411-2798.
2. Субботін С.О. Згортальна нейронна мережа як парадигма для реалізації технології глибокого навчання / С.О. Субботін, А.О. Олійник, О.Ю. Благодарьов. — ЗНТУ, 2015. — 1с.
3. Биков М.М. Метод нормалізації тривалості звучання паролних фраз для системи розпізнавання мовців / М.М. Биков, В.В. Ковтун, А. Раїмі ; Нац. Ун-т «Вінницький національний технічний університет». – Вінниця : Вид-во Нац. Ун-ту «Вінницький національний технічний університет», 2015. [Електронний ресурс]: Інформаційний портал. Режим доступу: <http://ir.lib.vntu.edu.ua/handle/123456789/3585?locale-attribute=en> .

Maksymov Oleksii O. – student of the group ЗАСІТ-17м, control computer system department, Vinnytsia national technical university.

Supervisor – Bisikalo Oleg V. – Dr. of Sci., professor, Dean of the Faculty of Computer Systems and Automatics, Vinnytsia national technical university.

Максимов Олексій Олексійович – студент групи ЗАКІТ-17м, кафедра комп'ютерних систем управління, Вінницький національний технічний університет.

Науковий керівник – Бісікало Олег Володимирович – доктор технічних наук, професор, декан факультету комп'ютерних систем та автоматики, Вінницький національний технічний університет.