UDC 004.8

NEURAL NETWORKS. COMPUTER VISUAL RECOGNITION

Gumen O.M., Doctor of Science,
gumens@ukr.net, ORCID: 0000-0003-3992-895X
Selina I.B.,
irinaselina2016@gmail.com, ORCID: 0000-0002-4010-3819
Miz D.S.,
fmf_ikg@ukr.net
National Technical University of Ukraine
«Igor Sikorsky Kyiv Polytechnic Institute» (Kyiv, Ukraine)

Neural networks are an important tool in modern technology, providing a wide range of applications in various fields, including computer vision, natural language processing, medicine, finance, and others. They are based on the idea of replicating the workings of the human brain and neural system, allowing them to perform complex tasks by using learning by example. Deep neural networks have become one of the most popular tools for developing automated solutions. These models turned out to be particularly useful in solving classification problems. Nevertheless, modern approaches to the formation of the network structure and learning require a significant investment of time. It is necessary to develop learning methods that could optimize the learning process in terms of the number of computational operations required. In addition, the method should have the property of adaptability, so that the learning results, as well as the obtained models, can be used for other tasks.

Segmentation, filtering, classification, scene reconstruction, object position estimation, object detection, video surveillance and many other diverse tasks include a wide field of computer vision. Computer vision is an important component of the development of artificial intelligence and intelligent information technologies. Image recognition is relevant in military affairs, when the correct identification of a military object can radically change the course of a battle or a local operation, finds its place in archeology and archival work, helps to reproduce historically valuable documentary finds, as well as when mapping the area and transmitting information from artificial Earth's satellites. It is desirable to give the reproduced image such properties that the perception of individual, most significant fragments of the image, or the image as a whole, will be convenient for the observer in view of the artistic or technological properties of this image. This paper examines the current state of neural network technologies, their operating principles and potential applications.

Key words: neural networks, artificial intelligence, computer visual recognition, natural language processing, technologies.

Formulation of the problem. In recent decades, neural networks in deep learning have opened up new possibilities in computer vision, where they are used for pattern recognition in videos and images. This covers a wide range of tasks, from object classification to face detection, motion and behavior analysis. In this study, we consider the main concepts, algorithms and applications of neural networks in computer vision, as well as the prospects for their development.

Analysis of recent research and publications. Deep neural networks have become one of the most popular tools for developing automated solutions [1]. These models turned out to be particularly useful in solving classification problems [2]. Nevertheless, modern approaches to the formation of the network structure and learning require a significant investment of time. Relearning of networks, if necessary, takes place on all examples of the learning data set [3]. These examples include those classes that the network successfully classifies (resulting in redundant work).

Formulation of the goals of the article. It is necessary to develop learning methods that could optimize the learning process in terms of the number of computational operations required. In addition, the method should have the property of adaptability so that the learning results (including the obtained models) can be used for other tasks.

Main part. After setting the network architecture, it is necessary to define its behavior. Learning a neural network consists in changing the value of the weights of the neurons so as to achieve the set goal of the task. Today, there are three main types of learning neural networks:

- supervised learning;
- unsupervised learning;
- semi-supervised learning.

In supervised learning, a set of learning data and corresponding outputs are sent to the network input. The challenge is that the network must find such a general rule that transforms inputs into outputs.

In unsupervised learning, the training sample is known, but the network needs to determine the internal relations between the data and independently form the outputs.

In mixed learning, the network receives an incomplete training set as input, some initial values are missing from it. And in this case, the network first learns on unlabeled data, and then, using this approximation, retrains on labeled data.

Computer vision is an extremely broad field that includes many diverse tasks such as segmentation, filtering, classification, scene reconstruction, object position estimation, object detection, video surveillance, and many others. Computer vision is an important component of the development of artificial intelligence and intelligent information technologies. Computer vision is used in dozens of industries, for example, when building "smart" stores, identifying customers using biological characteristics, automating agricultural processes

using drones, automatic inspection at factories, video surveillance, improving the quality of photo and video data (these are different architectures of neural network), automatic delivery of parcels by unmanned aerial vehicles.

Neural network architectures, namely convolutional neural networks (CNN), represent an interesting method for adaptive image processing and form a link between general feedforward neural networks and adaptive filters. Two-dimensional CNN are formed by one or more layers of two-dimensional filters with possible nonlinear activation functions and/or downsampling. CNN have the key properties of translational invariance and spatially local connections. Although the development of the CNN system for use is ongoing, the results support the notion that adaptive data-driven image processing techniques such as CNN are useful for image processing or other applications where the input arrays are large and spatially/temporally distributed.

Scheme of application of detectors (detector – an algorithm for finding special points) and descriptors (descriptor – a description of a special point that determines the features of its environment, is a numerical or binary vector of certain parameters) allows you to select a special point from the entire set of them in the image. This is necessary for compilation of key pairs of features belonging to the same object when comparing different images to solve the classification problem.

One of the most common classes of classification algorithms are the so-called bag-of-words (or bag-of-features, bag-of-key-points). The idea is borrowed from the problem of text classification, where a description is used in the form of histograms of occurrences of certain words from a pre-compiled dictionary. The main steps of such algorithms are described as follows:

Step 1 - identifying the key points of the image;

Step 2 - calculation of descriptors of local neighborhoods of special points;

Step 3 - clustering of descriptors of key points belonging to all objects of the training sample;

Step 4 - construction of a description of each image in the form of a normalized histogram of occurrences of "words" (for each cluster, the number of key points of a certain image attributed to it is calculated);

Step 5 - building a classifier that uses the image description calculated in step 4.

The other approach to solving the classification problem is to use models of objects consisting of parts (part-based models). Algorithms of this class take into account the relative location of various parts of the object. For example, when recognizing a face, is important to get recognized the relative location of the eyes, nose, mouth, hair, etc.

The main elements of models consisting of parts are:

- representation of separate parts of the object (usually descriptors are used for this);
- methods of learning this representation;

- description of connections between parts.

The procedure that ensures the conversion of electrical signals into an optical image is called visualization. However, the quality of the image is not always evaluated by the accuracy of reproduction of the primary image or the linearity of the transfer of parameter changes. For example, an image that was obtained at dusk or fog will be characterized by low contrast, indistinct outlines, pale color.

It is desirable to give the reproduced image such properties that the perception of individual, most significant fragments of the image, or the image as a whole, will be convenient for the observer in view of the artistic or technological properties of this image.

Conclusions. Image recognition is now very relevant in military affairs, when the correct identification of a military object can radically change the course of a battle or a local operation, finds its place in archeology and archival work, helps to reproduce historically valuable documentary finds, as well as in mapping the area and transferring information from artificial satellites of the Earth.

Література

1. Тимчишин Р.М., Волков О.Є., Господарчук О.Ю., Богачук Ю.П. Сучасні підходи до розв'язання задач комп'ютерного зору. 2018. № 6. С. 46-60. Режим доступу: http://usim.org.ua/arch/2018/6/8.pdf.
2. Браун М., Гідарі С.Ш. Згорткові нейронні мережі для обробки зображень: програма в Robot Vision.T.D. AI. 2003. LNAI 2903. С. 641-652. Режим доступу: https://link.springer.com/chapter/10.1007/978-3-540-24581-0_55.
3. Творошенко І.С. Цифрова обробка зображень: конспект лекцій. Харків: ХНУМГ ім. О.М. Бекетова, 2017. Режим доступу: https://core.ac.uk/download/pdf/83144138.pdf.

НЕЙРОМЕРЕЖІ. КОМП'ЮТЕРНЕ ЗОРОВЕ РОЗПІЗНАВАННЯ

Гумен О.М., Селіна І.В., Мізь Д.С.

Нейромережі - це важливий інструмент у сучасних технологіях, що забезпечує широкий спектр застосувань у різних сферах, включаючи комп'ютерне зорове розпізнавання, природну мову обробки, медицину, фінанси та інші. Вони базуються на ідеї відтворення роботи людського мозку та нейронної системи, що дозволяє їм виконувати складні завдання шляхом використання навчання на прикладах. Глибинні нейронні мережі стали одним з найпопулярніших інструментів розробки автоматизованих рішень. Дані моделі виявились особливо корисними при вирішенні задач

класифікації. Тим не менш, сучасні підходи до формування структури мережі та навчання вимагають суттєвих часових витрат. Необхідно сформувати методи навчання, які могли б оптимізувати процес навчання з точки зору кількості необхідних для виконання обчислювальних операцій. До того ж метод має володіти властивістю адаптивності, щоб результати навчання, а також отримані моделі можливо було використовувати для інших задач.

Сегментація, фільтрація, класифікація, реконструкція сцени, оцінка положення об'єкта, виявлення об'єктів, відеоспостереження та багато інших різнопланових задач включає в себе широка область комп'ютерного зору.

Комп'ютерне бачення є важливою складовою розвитку штучного інтелекту та інтелектуальних інформаційних технологій. Розпізнавання образів є актуальним у військовій справі, коли правильна ідентифікація військового об'єкта може кардинально змінити хід бою чи локальної операції, знаходить своє місце в археології та архівній справі, допомагає відтворити історично цінні документальні знахідки, а також при картографуванні місцевості і передачі інформації зі штучних супутників Землі. Відтвореному зображенню бажано надати таких властивостей, за яких сприйняття окремих, найбільш суттєвих фрагментів зображення, або зображення в цілому буде зручним для спостерігача з огляду на художні або технологічні властивості цього зображення. У цій роботі досліджується сучасний стан нейромережевих технологій, їх принципи роботи та потенційні застосування.

Ключові слова: нейромережі, штучний інтелект, комп'ютерне зорове розпізнавання, природна мова обробки, технології.

References

1. Tymchyshyn R.M., Volkov O.Ie., Hospodarchuk O.Iu., Bohachuk Yu.P. (2018). *Suchasni pidkhody do rozviazannia zadach kompiuternoho zoru*, 6. Retrieved from: http://usim.org.ua/arch/2018/6/8.pdf [in Ukrainian]
2. Brown M., Hydari S.Sh. (2003). *Convolutional Neural Networks for Image Processing: An Application in Robot Vision.T.D. AI 2003*, LNAI 2903. Retrieved from: https://link.springer.com/chapter/10.1007/978-3-540-245810_55
3. Tvoroshenko I.S. (2017). *Tsyfrova obrobka zobrazhen: konspekt lektsii.* Kharkiv: KhNUMH im. O.M. Beketova. Retrieved from: https://core.ac.uk/download/pdf/83144138.pdf [in Ukrainian]