COMPUTER VISION RETROSPECTIVE

Sardorbek Zokirov
Webster University in Tashkent, Adjunct Faculty

Abstract: *Significant growth in computer computing power and the emergence of new mathematical models and algorithms in recent years has made significant progress in the field of computer vision. New machine-based technologies enable you to create reliable solutions for complex visual image recognition tasks. Many of these solutions have practical applications in the market, and therefore attract more and more investment. This article is an overview of the basic approaches to solving computer vision problems, with a focus on deep machine learning and artificial neural networks. The article also provides an overview of market indicators in this area and shows that the computer vision market is developing rapidly and has serious prospects.*

Keywords: *Computer vision, machine learning, template search, artificial neural networks, deep learning.*

Computer vision (Computer Vision), and also sometimes machine vision (Machine Vision), is a scientific field engaged in research in the field of automatic fixation and various kinds of image processing (detection, tracking, identification) using a computer.

Some of the first attempts to create such systems were made in the 1960s, but due to the low power of computers at that time, this area was not sufficiently explored for a long time. Modern high-speed processors, cheap disk memory and high-quality cameras and high channel throughput make it possible to achieve significant results in this area, especially in connection with recent research in the field of machine learning and, in particular, neural networks.

The growth in the quality of computer vision technologies has allowed them to be actively used in various fields of industry and business. For example, many modern conveyors are equipped with automatic mechanisms for checking the quality of parts, correct labeling, alignment of parts and other tasks. An example of this approach is the microchip industry, where cameras automatically control the placement of silicon wafers and the layout of integrated circuits. It is possible to list a large number of areas of the digital economy in which modern computer vision is used:

SMART CITY SYSTEMS.

Infrastructure and transport systems with load control and balancing.

Automotive systems, including self-driving cars. Unmanned aerial vehicles, including drones. Pharmaceuticals and medicine. Video analytics.

Recognition of people and objects. Identification of persons. Assessment of the scene, time, place.

The main purpose of computer vision is to obtain useful information from an image or a series of images. Tasks can be such as:

■ Calibration of optical systems, setting up and synchronizing camera settings,

■ detection of object motion (Motion Tracking),

■ Object recognition tasks,

■ Scene reconstruction tasks (usually 3D scenes from one or more 2D frames),

■ tasks of comparing images and identifying changes.

Artificial neural networks: emergence, development and current state

A standard neural network consists of many simple connected processors called neurons, each of which produces a sequence of activations (vectors of real numbers). Entrance neurons (neurons of the first, input, layer) are activated by sensors that perceive the environment, the rest are activated by weighted connections with other neurons that were activated earlier.

Neurons can influence the environment by triggering actions. Learning (or trust assignment) consists in finding such weights with which the neural network achieves the desired result, for example, recognition of a reference sample. Depending on the task and the type of neuronal communication, this behavior may require a chain of steps, each of which non-linearly transforms the activation of previous steps. Deep learning is the training of networks with many such steps.

Shallow models like neural networks have been around for decades. The method is called the error back propagation algorithm. However, it was impractical to train deep neural networks using this method due to the lack of available computing resources. Such problems became the objects of research only in the 1990s.

Back then, teaching without a teacher was mainly used, but teaching methods with a teacher were improved.

Direct (acyclic) and recurrent neural networks have been successfully used in a variety of tasks. Recurrent neural networks are the deepest, in principle they are universal computing systems and can, in theory, process and store any patterns. Unlike classical methods, recurrent neural networks can be trained to create programs that effectively mix sequential and parallel computing.

The topology of a neural network may change over time. At a certain point in time, it can be written as a finite set of vertices (neurons) N = {sch, sch, ...} and a finite set H with N x N directed weighted edges (connections) between vertices. The graph of a direct neural network is acyclic, the graph of a recurrent network is cyclic. The first (input layer) is a set of input neurons. In acyclic neural networks, the ^th layer ^ 6 N k > 1) is the set of all vertices u 6 N such that there is a path of length (k - 1) between one of the input neurons and one of the neurons w

The behavior of a neural network is determined by a set of real parameters (weights) (£ = 1...n). For now, let's focus on one era of information dissemination,

not paying attention to learning and changing weights. That is, let's describe what happens when a trained neural network is running.

The input neuron receives some kind of real value as input. This value is propagated, that is, it is copied to the input to all neurons connected to the original one, while being multiplied by the weight of the edge. Each of these neurons adds up all the incoming values, and applies a nonlinear activation function to the sum (for example, f(x) = max(0, x)) and distributes the resulting value further over the network. The activation value of the neurons of the last layer is the output value of the neural network. There are a large number of additional operations used in this procedure (for example, batch normalization: after the selected layer, all neurons change linearly so that the average value of all neurons is zero and the variance is one), but they are optional and applied depending on the task.

In teaching with a teacher, the vector of output values xt of a neural network can be compared with the vector set by the teacher (manually, another network, or any other way) dt, while the difference between them can be considered an error value, for example et = % (xt - dt)2. The task of training is to minimize this error using the gradient descent method.

Simple and complex cells were found in the retina of the cat's eye. The labels signaled in response to certain characteristics of the visual image, for example, the vertical or horizontal location of the boundaries of objects. Complex cells reacted to more complex patterns than simple cells. In the future, these observations became the basis for convolutional neural networks, an architecture that has so far been successfully used in solving pattern recognition problems.

REFERENCES:

1. Дэвид Форсайт, Жан Понс. Компьютерное зрение. Современный подход = Computer Vision: A Modern Approach. — М.: «Вильямс», 2004.

2. Л. Шапиро, Дж. Стокман. Компьютерное зрение = Computer Vision.— М.: Бином. Лаборатория знаний, 2006.

3. Goodfellow, Y. Bengio, A. Courville. "Deep Learning Book", MIT Press,2016 http :// deeplearningbook.org.

4. M. Nielsen, "Neural Networks and Deep Learning", DetermintationPress, 2015 http://neuralnetworksanddeeplearning. com.

5. W. Dai, C. Dai, S. Qu, J. Li, S. Das "Very Deep Convolutional Neural Networks for Raw Waveforms", ICASSP, 2017