Computer vision is an important field of science and technology. Currently, there is a significant increase in the use of computer vision techniques in many areas, in particular in medicine, army, robotics, monitoring and surveillance systems, entertainment and film industry. Vision systems will play a key role in autonomous vehicles, aircrafts, humanoid robots and artificial reality systems. In significant number of practical applications, quality and effectiveness of the perception are influential factors on the possibilities of the entire system. Moreover, without significant progress in the field of sciences related to learning of the recognition systems, and in particular without significant development of decision-making systems, which operate without human intervention using data delivered by sensors, it will not be possible to deploy service robots on a larger scale, or autonomous vehicles. The main obstacle to the dissemination of these technologies, as well as many related technologies in which image sensors play an important role, is that currently used algorithms for detection, classification and recognition require a significant number of examples to build a model or to learn a neural network. The immanent feature of currently available algorithms is that they require hundreds or even thousands of times more sample images at the learning stage compared to what a human needs. What's more, most of the currently available algorithms are not resistant to noise, in particular to all kinds of artifacts and data distortions. Another significant limitation of modern algorithms is that they lack the ability to generalize. Consequently, it is much easier to build a system that will win in human in checkers and even chess, rather than build an algorithm that would allow a humanoid robot to capture an unknown object on the basis of camera images.

Currently, deep neural networks are widely used in many areas of science. In vision systems, deep neural networks that include convolutional networks and classical networks with fully interconnected neurons are quite commonly used. Convolutionary layers are responsible for separating the features with the highest discriminative power, while fully connected layers are responsible for classifying the features extracted in such a way. Deep neural networks consist of many layers that in subsequent layers group features on maps generated by previous layers into more general representations of object features. The costs and workload associated with the preparation of collections of labeled data for the purposes of learning deep convolutional networks, as well as relevant deep networks are significant. These costs entail the need to develop unsupervised learning methods that are quite often used in the first stages of learning.

Currently, there is a need to create new or optimize already existing algorithms dedicated to building and learning deep networks. In particular, there is a need for new and more effective solutions for the needs of broadly understood computer vision. Scientific research that will be carried out in this project concerns the development of models and learning of deep structures composed of heterogeneous neural networks and blocks of image analysis/rendering. The developed solutions will enable more effective use of non-labeled data, and in particular will result in effective learning mechanisms of neural networks, which, although they will be learned on significantly smaller data sets, will have a comparable discriminating power. Thanks to the developed solutions, the proposed algorithms will enable recognizing not only objects on the basis of examples that were used for learning, but also objects never previously seen, with support of discovered relationships and dependencies of a more general nature. The result of this project will be not only solutions and studies that will give answers to many questions, but also algorithms that will be evaluated, as well as specific solutions for real-time computer and robot vision.