The perception of robots is mainly based on the modeling of geometric structures in the environment. Wheeled mobile robots, walking robots and manipulators use RGB cameras, depth sensors, and laser scanners to determine the shape of obstacles in the environment. For this purpose, occupancy maps, raster maps, voxel maps are built which describe occupancy of the space and shapes of obstacles. Knowing the shape of obstacles, robots plan their movements and avoid collisions with the environment, but do not know the meaning and relationship between the objects in the environment.

In the traditional approach to robot perception, every object is treated as an obstacle. In reality, some objects can be moved by the robot, the door can be opened to allow passage. The first use of deep neural networks in robotics allowed the classification of objects in color images. Thanks to this, the robots can select objects to grasp, avoid collisions with people, and improve their localization. It is also possible to classify the terrain, which provides traversability cost during motion planning of wheeled and walking robots.

The next step in the development of perception systems and, at the same time, the subject of this project is to inference about the properties and meaning of objects. An example of a scenario is the use of neural networks to estimate the properties of objects such as doors, drawers, switches, etc. In this scenario, a robot that uses a two-dimensional representation of the environment (RGB image) and a depth image should conclude on the potential motion of the articulated objects, its kinematic limitations, and the state. Another example is the reconstruction of 3D objects. Robots, unlike humans, are not able to reconstruct 3D objects using a single pair of RGB-D images. A modern robot perception system should also reconstruct surfaces that are unseen because they are occluded by other objects or are on the other side of the object. Such properties of perception systems can be obtained by using artificial neural networks and aggregating knowledge about the robot environment and object properties with the application of learning mechanisms.

With such a perception system, the robots will be able to better plan their motion and interaction with the environment based on single RGB-D pair images, without the need for time-consuming scanning and building a model of the environment. The aim of the project is also to find an answer to the question of how robots should represent the environment and how robots "understand" the meaning of surrounding objects, their shape, and kinematic relations.

The project aims to prove that robots can infer about object properties from a limited number of RGB-D images without the need for careful space scanning and direct interaction (touch, object manipulation, etc.). Full models of manipulated objects can be reconstructed based on one pair of RGB-D images and the use of a neural network that stores information about the shape of the reconstructed objects. We are going to reconstruct everyday household objects, which the robot grabs and manipulates them, as well as elements of room equipment such as furniture, doors, walls but also whole rooms. Algorithms created during the project will allow predicting the state of the robot's environment before the execution of the motion. These methods will be integrated with neural methods that predict the state and limitations of the robot during the planning of its motion in the environment and make the robot more aware of the environment and its limitations.