Implementation of Unmanned Control of Wheeled Robots

M.S. Lyubimov[1], V.I. Lushkov [1], A.A. Azarchenkov[1]
max32@inbox.ru|lyshkoff2017@yandex.ru| azarchenkovaa@yandex.ru
[1]Braynsk State Technical University, Bryansk, Russia

This paper presents an approach to the unmanned control of a wheeled robot, which includes recognition of road infrastructure objects, recognition of continuous and intermittent road markings, generation of control signals. Recognition of road infrastructure objects is carried out using a neural network that generates a segmented image. After that, the segmented image is identified with the found objects, including the roadway, which is used by the road marking recognition subsystem searching for continuous and intermittent lines using the computer vision library. On the basis of the information received from the considered subsystems control commands are generated indicating the direction of movement and speed. The algorithm was developed on a 1:18 scale model of the city infrastructure, where a wheeled robot simulated as a car.

Keywords: *neural network, object localization, image segmentation, Pyramid Scene Parsing Network, wheeled robot, road marking recognition*.

1. Introduction

At present there are active developments in the field of unmanned vehicles. Some companies are approaching the last, fifth level of autonomy, in which autonomous movement of the car without the driver becomes fully possible.

The task of providing unmanned motion of a wheeled robot has been described in such projects as:

- Duckietown [9] - is a project aimed at the study of robotics and artificial intelligence, using available robotic platforms and training programs for all levels of education. Duckibot is a two-wheeled platform which has raspberry pi mounted with a camera connected to it, an engine driver, two engines and a powerbank. It is controlled remotely via wi-fi. A laptop with Ubuntu system is used as a computing node. The ROS framework was used to write software modules. The robot in question is able to travel along the marking lines on the city model. Marking is searched by identifying contours and filtering by color value. The stop line is processed in the same way. Identification of road signs is performed using a neural high-precision network. There is also processing of all signals and reaction to detected objects.
- Formula pi [10] is a project which deals with the development of algorithms for unmanned motion of wheeled robots. The movement is carried out along the race track, where the car lanes are highlighted in different colors. The algorithms are executed on Raspberry pi platform, which narrows the range of approaches possible for application. The camera is used as a sensor. Algorithm principles are different, which is the main idea of the project, but based on the recognition of marking lines.

One of the key tasks without which it is impossible to create an unmanned vehicle in a dynamic environment is the task of recognizing road network objects. There are different approaches to this task: from the recognition of individual elements, such as road signs or traffic lights [7], to more comprehensive approaches [8], allowing to detect several elements. Recognition can be made on the basis of information coming from different sensors. The most common sensors are cameras that return images and lidars that return a point cloud around them. The use of cameras is a more traditional method that simulates the driver's vision. In the field of computer vision, object recognition is one of the main tasks, for which both standard algorithms of computer vision [4] and different architectures of neural systems [1] are used. The method described in this article is based on the use of a neural network [6], which performs the task of image segmentation.

The developed object recognition system was tested on a model containing several types of traffic lights, road signs, pedestrians, cars, roadway with marking.

2. Recognition of Road Infrastructure Objects

The basis of object detection is a neural network that receives the input image from the camera and returns a segmented image (Fig.1). Segmentation is the process of generating an image that contains homogeneous areas of the same color, where each color corresponds to a specific object. Recognition of objects based on the segmented image allows to solve several tasks necessary for the unmanned movement of the vehicle: on the one hand, it allows to detect objects of the road network, on the other hand, the problem of roadway recognition is solved.

Fig. 1 Example of image segmentation

The neural network used is based on Pyramid Scene Parsing Network architecture[3]. The main components of the network are the pre-trained Residual Network (Resnet) [5] and Pyramid Pooling Module.

Reset model is a convolutional model. Its distinctive feature is the availability of a residual learning structure which allows to increase the network complexity without loss of accuracy. At the output of the convolutional network, a feature map is generated, which enters the pooling layers. This part of the network is called

Pyramid Pooling Module because it consists of several levels (Fig. 2):

1) red: this is the largest level that performs the main pooling for each feature map, resulting in a single output value;
2) orange: this is the second level that divides the feature map into 2 × 2 subregions and then performs the average combine value for each subregion.
3) blue: this is the third level that divides the feature map into 3 × 3 subregions and then performs the average combine value for each subregion.
4) green: this is the level that divides the feature map into 6 × 6 subregions and then performs the combine value for each subregion

After that upsampling of layers and their combining are made.

The resulting model is able to recognize such classes of objects as road signs, traffic lights, pedestrians, vehicles, roadway, and pedestrian crossings.

It should be noted that it was decided to recognize road signs and traffic lights in two stages in order to increase the accuracy of recognition. At the beginning, the object is localized in the image using the described neural network, then the classification network is used to determine the type of a road sign and a series of techniques, which are based on the definition of color in the image, to determine the signal of the traffic light.

3. Road Marking Recognition

Road markings is recognized using OpenCv computer vision library. Processing of the input image, getting rid of noise, and smoothing are performed. After that, marking lines in the image are detected together with the road segment obtained from the road infrastructure recognition module.

The search for a continuous marking is done by combining segments into strips detected by means of Hough transformation.

Intermittent marking is detected by analyzing the contours of the image and selecting the most similar of the distinctive features such as the angular coefficient and aspect ratio.

The last stage is the construction of virtual lanes that stabilize driving along straight sections of the road, as well as the construction of the trajectory and the calculation of steering angle (Fig. 3).

4. Module of Generating Control Signals

Data detected by sensors and cameras are used to control a wheeled robot for generating control signals which relate to three types: speed increase or decrease, change of steering angle, direction change of the wheeled robot motion.

Each object, which can include road signs, pedestrians, vehicles, traffic lights, range measuring systems, has its own set of states: do nothing, stop, move forward, increase or decrease the speed.

Then, all incoming data are analyzed and each object is assigned with the state, depending on the distance defined to this object and other conditions.

Data sent to the module of generating control signals are divided by priority, having the following sequence:

1) data from the wheel robot controller.
2) data from visual data analysis system.
3) data from the navigation module.
4) data from the marking module.

At the end the control signal is generated, which is sent to the wheeled robot (Fig. 4).

Once the traffic light has been identified in the image, the data about it comes to the module of controlling a wheeled robot based on visual data, and then there is a stop or continuation of the robot movement, depending on the signal of the traffic light.

If the road sign is detected the following search of signs is made: major road, yield sign, crosswalk, slippery road, no entry, stop, canceling former signs, do not pass. After that, the command is given to the wheeled robot to stop, continue moving, increase or decrease the speed.

When a vehicle is detected, stop will be made, and if there is a corresponding sign, the vehicle will be given the way.

If a pedestrian is detected, the wheeled robot will stop in front of him and wait for him to disappear from the frame.

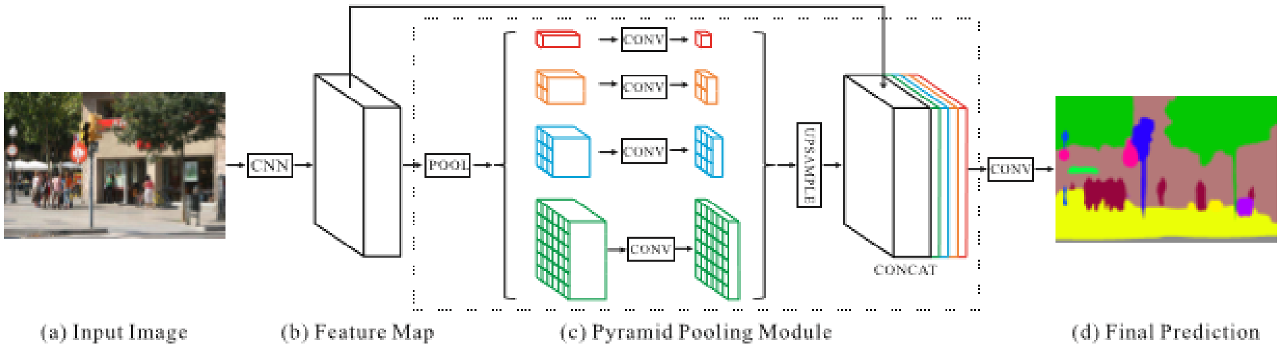

Fig. 2 Pyramid Science Network

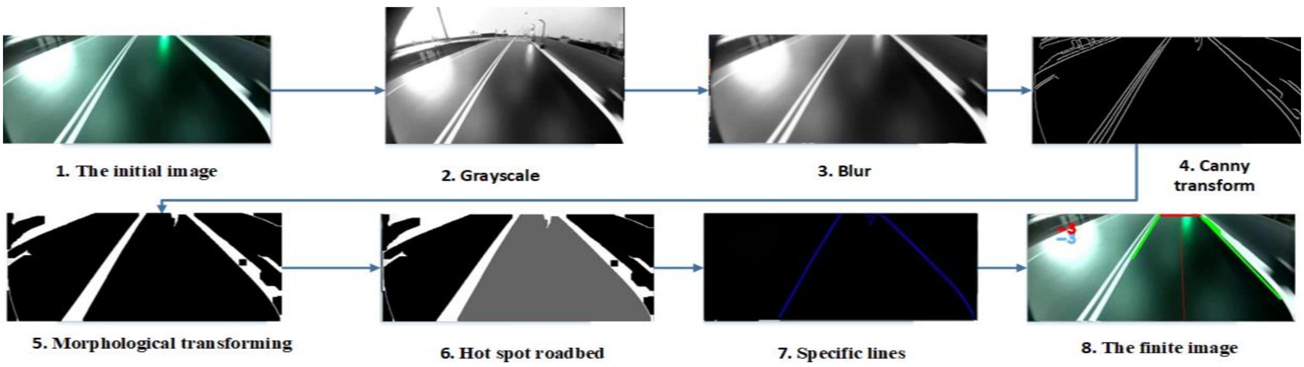

Fig. 3 Stages of marking recognition

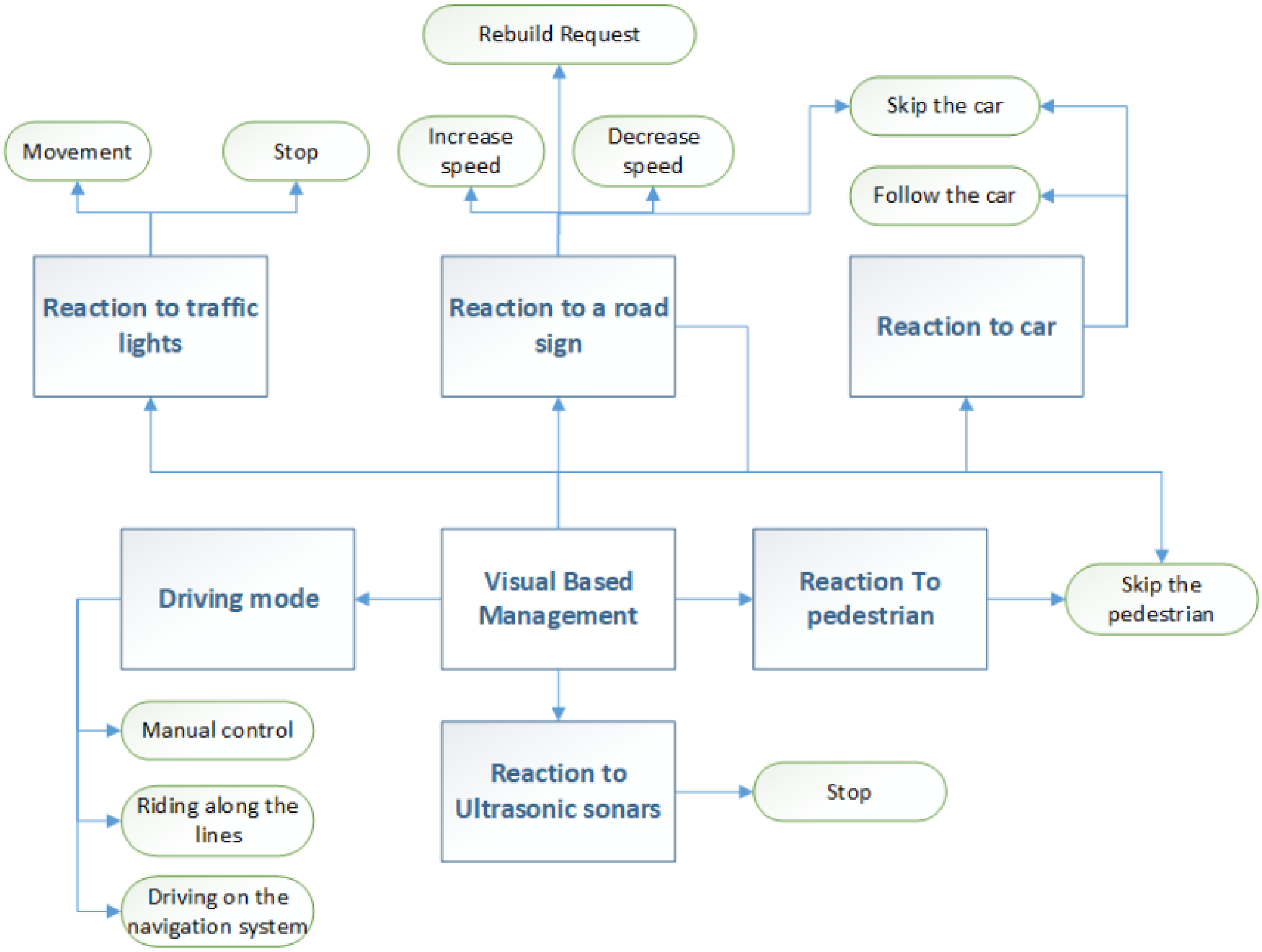

Fig. 4 Stages of marking recognition

5. Experiments

The described approach is used for a wheeled robot's moving through the model layout of the urban infrastructure (Fig. 5.). The layout has a size of 6x4 and is equipped with a roadway with markings, road signs, other cars, pedestrians, traffic lights, including traffic lights with an additional section. The wheeled robot is shown in figure 6. The peculiarity of the architecture is that all calculations take place at a remote site. This makes possible to work with modules that require higher performance and make the layout near real vehicles.

.

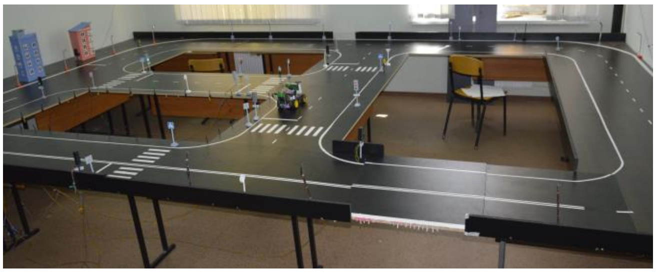

Fig. 5 Road infrastructure model layout

The launch of the neural network was carried out on the NVIDIA GTX 1080 graphics card. The average generation time of a segmented image is 0.063 seconds, which provides sufficient speed to move around the layout. For testing, the resulting neural network and object allocation subsystem were integrated into the main project produced by using the ROS framework. The first node subscribes to the camera image, generates a segmented image, and publishes it.

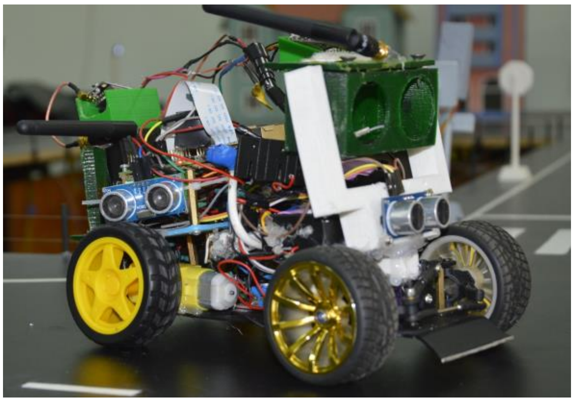

Рис. 6 Wheeled robot

The second node subscribes to the ordinary and segmented image, selects objects and publishes the following messages: roadbed, road signs, which are sent to be classified, traffic light that is sent to the module of signal determination, a message containing pedestrians and the automobile, which is sent to the merge module of objects.

The found objects are added to the database, and after that they are sent to the control signal generation node, where data analysis and creation of control commands for the wheeled robot take place.

6. Conclusion

Within the work an approach providing the unmanned control of a wheeled robot was developed and tested. Within the framework a subsystem localizing the main objects of road infrastructure such as roadbed, pedestrian crossings, cars, traffic lights, road signs, was prepared. Subsystems of road signs classification, traffic signals, road marking recognition, control signals generations were also developed. The results were tested on the model layout simulating an urban environment.

7. References

[1] A. Buyval, M. Lyubimov, A. Gabdullin "Road sign detection and localization based on camera and lidar data" //The 11th International Conference on Machine Vision (ICMV 2018) — Munich, Germany 2018

[2] Cordts M., Omran M., Ramos S. ,Rehfeld T. ,Enzweiler M. , Benenson R., Franke U., Roth S., Schiele B. “The Cityscapes Dataset for Semantic Urban Scene Understanding”/ M. Coordts. 2016г.

[3] Hengshuang Z., Jianping S., Xiaojuan Q., Xiaogang W., Jiaya J.” Pyramid Scene Parsing Network”/ Z. Hengshuang. CVPR 2017г.

[4] Md. S. Hossain, Z. Hyder “Traffic Road Sign Detection and Recognition for Automotive Vehicles”, International Journal of Computer Applications (2015)

[5] Kaiming H, Xiangyu Z ,Shaoqing R ,Jian S “Deep Residual Learning for Image Recognition” / H. Kaiming, Z. Xiangyu, R. Shaoqing, S Jian – Microsoft Research 2015г

[6] J. Long, E. Shelhamer, T. Darrell "Fully Convolutional Networks for Semantic Segmentation" // CVPR 2015

[7] A. Møgelmose, M. M. Trivedi, T. B. Moeslund, "Vision based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey," IEEE Transactions on Intelligent Transportation Systems, (2012).ё

[8] A. D. Pon, O. Andrienko, A. Harakeh, S. L. Waslander "A Hierarchical Deep Architecture and Mini-Batch Selection Method For Joint Traffic Sign and Light Detection" IEEE 15th Conference on Computer and Robot Vision (2018)

[9] Duckietown [Electronic resource] – URL: https://www.duckietown.org, available.

[10] Formula Pi [Electronic resource]. – URL: https://www.formulapi.com, available.