Autonomous Car Computer Vision

Filatov, D.M.; Devyatkin, A.V.; Serykh, E.V.

Abstract

In the paper active driver assistant subsystem of sign recognition design and evaluation are considered. Multiple approaches (color operators and deep neural networks) in cases of hardware limitations with environment simplifications are utilized. Simple color and morphology operations allowed using designed system on an embedded system based on ARM 32bit microprocessor. The growth of computation power from small embedded microcomputer to module system based on Intel x86 64 bit CPU with ARM 128bit Pascal GPU allowed to change the utilized approach and its complexity to use deep learning neural network with truncated structure of commonly used U-net. Last approach is widely used in the world competition of traffic signs recognition systems. The achieved results and experience are connected to autonomous cars competitions.

1 INTRODUCTION

The recent competitions in the car active driver assistant (ADAS) represent the requirement of stable traffic sign recognition subsystem. First, the main requirement to the subsystem is the correct and accurate sign recognition for further system processing. Nowadays, traffic signs represent and describe the rules of car behavior on the road. Starting from second and third levels of automated driving system the car starts monitoring the environment and decides what to do in cases of different traffic signs.
Developments and research in the field of autonomous vehicles are currently very active. As an example, the route construction of the vehicle in operation can be studied [1], the development of a decision logic module in the behavior of the vehicle [2], the use of the tech-nology of autonomous vehicles to help blind people to navigate [3], moreover, the development of a multi-functional system for an autonomous machine based on heterogeneous sensors for distance measurement [4].
The design of traffic sign recognition (TSR) system is based on processing of the information about environment. The most used information nowadays is visual information taken from visual sensors like cameras. The received information is represented as set of colored points that carry required information and the base of algorithm is to extract required information about traffic signs for further processing. The mostly used algorithms are considered form the field of computer vision and the approaches [5-9] have high potential for development but the accuracy is not high and changes of environment can cause errors in recognition.
The paper represents the steps of the system growth and embedding experience of the project design that includes traffic sign recognition task.

2 PROTOTYPE DESCRIPTION

All of the development is designed on the small radio control car model (Fig. 1). The sizes are represented in the table (Fig. 2). HPIWR8 Flux is a four-wheel drive rally car with FordFiestaH body.F. H. V. KenBlock and brushless motor (scale 1:8). Radio-controlled car model is able to develop a speed of over 97 km/h, the chassis has reliable differentials with bevel gears and ball bearings in all nodes of the four-wheel drive transmission. Rally suspension geometry with specially designed racks shock mount, heavy-duty shock absorbers and thicker anti-roll bars guarantees control and scale RallyCross tyres on Speedline wheels provide excellent grip.

Fig. 1 Rally car model HPIWR8 Flux

Parameter	Value, mm
Length	485
Width	227
Height	172
Wheel base	300
Wheel diameter	80

Fig. 2 Rally car model parameters

3 SIMPLE COLOR APPROACH

The design of the system for the first competitions was based on single board computer Raspberry Pi 2 Model B (RPi) (Fig. 3). The mentioned microcomputer has 900MHz quad-core ARM Cortex-A7 CPU and 1 GB RAM of memory. The main advantage of the microcomputer is its size and number of interfaces on it. On the other side, the computation power of the microcomputer is quite pure, that’s why the first approach for traffic signs recognition task was based on main CPU non-critical functions.

Fig. 3 Raspberry Pi 2B picture

The set of competition signs is represented in the figure 4.

Fig. 4 Set of competition signs

The basics of the system were sign extraction region with morphological operators and feature extraction with sign classification. Library OpenCV was used for all matrices processing.

The morphological operators processing example is represented in the figure 5. The preprocessing of operators was Gauss filtering. The morphological operators were “erode” and “dilate” with color filtering for sign localization.

Fig. 5 Morphological operators processing

Further processing was based on contour selection that describe the sign region with geometrical figure. An example of contour selection is in the figure 6.

After ROI (region of interest) was extracted feature extraction with classification should be done. The base of current feature extraction was based on construction of 20 vertical and 20 horizontal parameters from gray-scaled image with mean values of HSV components. As a result, 43 features were inputs of self-written neural network. The type of the network is perceptron and structure is in the figure 7.

Fig. 6 Contour selection on picture

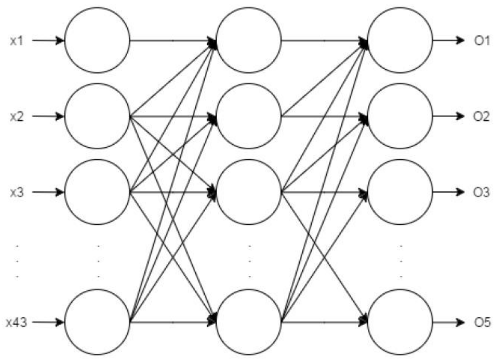

Fig. 7 Neural network structure

The structure of the neural network is (43-60-5). Threshold value of correct sign recognition is 0.98. An activation function of the neuron is sigmoid.

The speed of the full system processing was 15 FPS. The true positive recognition was 87% for competition set of signs.

4 DEEP LEARNING APPROACH

The described approach was developed for special set of signs and it is extension for wider set of signs has limitations. The requirement is to design a system with capability of adaptive feature extraction.

As used in the previous system, machine learning approach that was neural network requires learning set with new signs extend list of recognition set. Otherwise, extraction algorithm is only for round sign with special colors.

The deep learning approach with convolutional layers has great utilization [10-12]. The developed system is based on Intel NUC x86 CPU architecture with 64-bit address bus width. The additional unit is GPU NVIDIA Jetson TX2 with Pascal architecture with 128-bit address bus width.

The communication of CPU and GPU is realized with Ethernet TCP/IP stack and ROS communication protocol. The main bus scheme of system with two subsystems is represented in the figure 8. The basic libraries for image processing and machine learning methods utilization were OpenCV, Keras and SciPy.

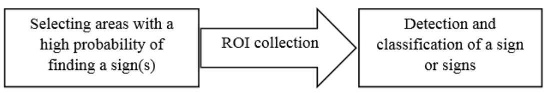

Fig. 8 Neural network bus system

The first subsystem is deep learning neural network that realizes ROI selection based on heatmap creation. The basis of the network is the principle of special features allocation on the image and further transformation into a matrix representa-tion, which is often the scaling of the original image. The system uses the structure “encoder-decoder” that can be constructed with Fully Convolutional Neural Network (FCNN). The task of image segmentation was taken as a source of ideas.

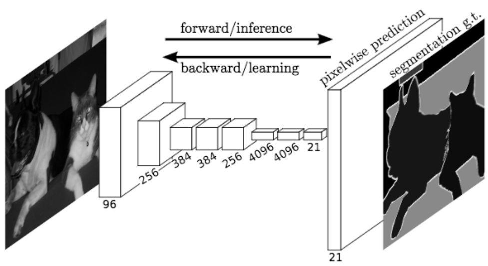

Fig. 9 Example of FCNN structure

The main rates for evaluation were Sørensen–Dice coefficient (1) and Jaccard coefficient (2) evaluation functions:

$$S(A,B) = \frac{|A \cap B|}{|A| + |B|} \quad (1)$$

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

An example of the FCNN processing is in the figure 10.

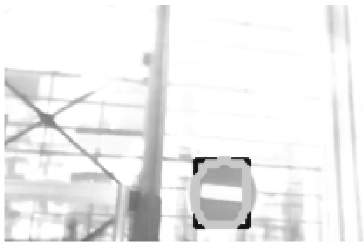

Fig. 10 FCNN sign ROI selection processing

The network training was performed with stochastic gradient descent (SGD) with varying learning rate.

Next classification task was performed. The base was taken as combination of Histogram of Oriented Gradients (HOG) for feature extraction and Linear Support Vector Machine (LSVM) for feature classification.

The principle of HOG is to represent the image as a vector histogram built on the image gradient. The method maintains the invariance of the movement of the object in the image. The principle of LSVM is the division of the feature space based on the principle of maximum boundaries (figure 11).

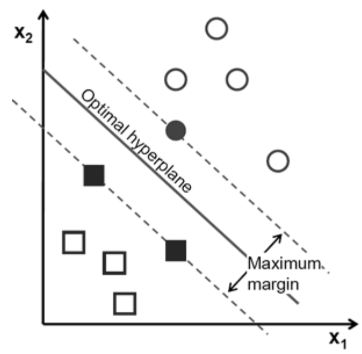

Fig. 11 LSVM processing example

The performance of the network was estimated about 50 FPS with GPU utilization and 14 FPS only with CPU support. That result allows to increase the complexity of the network for learning to critical cases like high brightness and high image distortions. The accuracy was performed about 87% according to F-score based on the confusion matrix.

5 CONCLUSION

The described approaches represent the consistency of introduced assumptions. The designed RPi system allows to utilize TSR system for low power and size critical plants. The question of complexity to increase the accuracy or number of signs requires the system with learning capability. It represents the high potential in image transformations. The only challenging question is the learning samples development.

6 REFERENCES

[1] Gaku Takano, Makoto Obayashi, Keisuke Uto, «Posture-to-posture trajectory planning for an autono-mous car using triple clothoid segments», 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, DOI: 10.1109/CDC.2017.8263682

[2] Bingqing Xu, Qin Li, «A bounded multi-dimensional modal logic for autonomous cars based on local traf-fic and estimation», 2017 International Symposium on Theoretical Aspects of Software Engineering (TASE), Sophia Antipolis, France, France, DOI: 10.1109/TASE.2017.8285637

[3] Manuel Martinez, Alina Roitberg, Daniel Koester, Rainer Stiefelhagen, Boris Schauerte, «Using Technology Developed for Autonomous Cars to Help Navigate Blind People», 2017 IEEE International Conference on Computer Vision Workshop (ICCVW), Venice, Italy, DOI: 10.1109/ICCVW.2017.169

[4] Ayesha Iqbal, Syed Shaheryar Ahmed, Muhammad Danish Tauqeer, Ali Sultan, Syed Yasir Abbas, «De-sign of multifunctional autonomous car using ultrasonic and infrared sensors», 2017 International Sympo-sium on Wireless Systems and Networks (ISWSN), Lahore, Pakistan, DOI: 10.1109/ISWSN.2017.8250023

[5] M. Shneier, “Road sigh detection and recognition”, Proc. IEEE Computer Society Int. Conf. on Computer Vision and Pattern Recognition, pp. 215 – 222, 2005.

[6] A. Ruta, F. Porikli, Y. Li, S. Watanabe, H. Kage, K. Sumi, “A New Approach for In-Vehicle Camera Traffic Sign Detection and Recognition”, IAPR Conference on Machine Vision Applications (MVA), Session 15: Machine Vision for Transportation, 2005.

[7] R. Belaroussi, P. Foucher, J.P. Tarel, B. Soheilian, P. Charbonnier, N. Paparoditis, “Road Sign Detection in Images”, A Case Study, 20th International Conference on Pattern Recognition (ICPR), pp. 484 – 488, 2010.

[8] S. Houben, J. Stallkamp, J.Salmen, M. Schlipsing, C. Igel, “Detection of Traffic Signs in Real-World Imag-es: The German Traffic Sign Detection Benchmark”, International Joint Conference on Neural Networks, 2013.

[9] J.Stallkamp, M.Schlipsing, J.Salmen, C.Igel, “Man vs. computer: Benchmarking machine learning algo-rithms for traffic sign recognition”, Neural Networks, Vol. 32, pp. 323 – 332, 2012

[10] Prashengit Dhar, Md. Zainal Abedin, Tonoy Biswas, Anish Datta, “Traffic sign detection — A new approach and recognition using convolution neural network”, 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, Bangladesh, DOI: 10.1109/R10-HTC.2017.8288988

[11] Tong Guofeng; Chen Huairong; Li Yong; Zheng Kai, “Traffic sign recognition based on SVM and convolutional neural network”, 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), Siem Reap, Cambodia, DOI: 10.1109/ICIEA.2017.8283178

[12] Kaoutar Sefrioui Boujemaa; Ismail Berrada; Afaf Bouhoute; Karim Boubouh, “Traffic sign recognition using convolutional neural networks”, 2017 International Conference on Wireless Networks and Mobile Communications (WINCOM), Rabat, Morocco, DOI: 10.1109/WINCOM.2017.8238205

Authors: Alexey Devyatkin; Denis Filatov; Elena Serykh

University: Saint Petersburg Electrotechnical University "LETI"

Department: Automatic Control Systems

E-Mail: avdevyatkin@etu.ru

dmfilatov@etu.ru,

evserykh@etu.ru