Traffic sign detection and problems in the field of computer vision

Marina Lukashevich
Department of Electronic Computing Machines
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
lukashevich@bsuir.by

***Abstract*—Object detection is a typical task of computer vision. This paper presents some results of implementation of the traffic sign recognition. We use R-CNN for traffic sign detection system. We focus on speed limit superclasses of traffic sign. R-CNN deep learning detector is a simple and suitable model for the traffic sign recognition. This approach combines multiple low-level image features with high-level context from object detectors and scene classifiers. Despite the existing advances in computer vision, the article considers the problems that exist and which need to be solved in the future in the field of computer vision system design.**

***Keywords*—object detection, deep learning, R-CNN, intelligent system, computer vision**

I. Introduction

In recent years image processing and analysis, pattern detection are the most exciting and fastest-growing research areas in the computer vision. Recent computer vision technologies and algorithms are support efficient semantic image segmentation and classifications. Intelligent driver assistance systems are systems to help the driver in the driving process. Traffic sign recognition is a technology by which a vehicle is able to recognize the traffic signs put on the road. This is part of the Intelligent driver assistance systems [1]. The goal of our research is to detect traffic signs into speed limit superclasses as shown in Fig. 1.

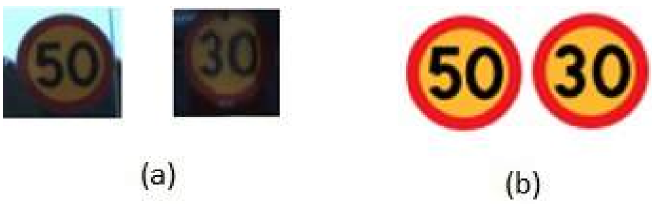

Figure 1. Examples of real-life traffic signs superclasses (a) and there synthetic examples (b).

Object detection is the task of finding the different objects in an image and classifying them. In our research, instead of developing of traffic sign detection system by applying R-CNN detector we try to formulate problems and goals that exist in the field of computer vision. Computer vision is one of the fields of Artificial Intelligence that has grown the most in the last 15 years. But we can speak only about private decisions of individual narrow tasks of computer vision.

This paper is structured as follows. Session II-IV describes image processing stages and used algorithm. In session V we formulate problems and goals in the field of computer vision.

II. Related research studies

A comprehensive review on the recent achievements of traffic sign recognition was presented in [2]. Many different approaches have been used for traffic sign recognition. Typical sign detection algorithms consist of three stages: segmentation, detection and classification, Fig. 2.

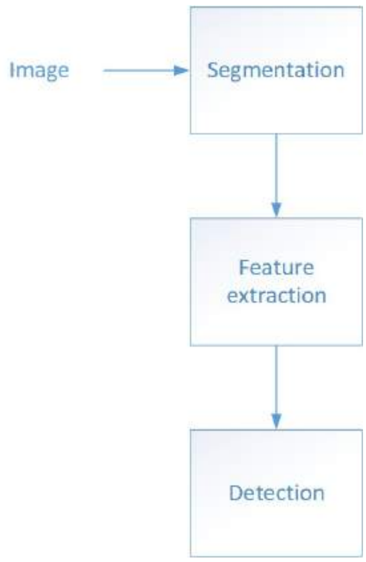

Figure 2. Typical sign detection algorithms

Traditional approach for segmentation is to threshold in chosen color space (HSV, HSI or CBH) [3-5], in order to obtain a binary image, Fig. 3.

Feature extraction is the second part of this process. In [6] authors have calculated Fourier descriptors. Other approach is using various HOG-features or edges [2]. The last stage is detection and classification using suitable classifier (SVM, neural network, Cascaded classifier, Fuzzy templates, Hough shape detection, etc.) [2], Fig. 4.

Figure 3. Example of segmentation process

Figure 4. Example of traffic sign detection

III. R-CNN OBJECT RECOGNITION

This typical scheme does not always provide the required accuracy of the detection. Traditional machine learning makes it possible to obtain good results with small datasets. It is possible to train a model quickly. But there is an accuracy plateaus in this case. We need to try different features and classifiers to achieve best results.

Recently, deep learning have significantly improved image classification and object detection accuracy [7]. Traditional neural networks contain only 2 or 3 layers, while deep networks can have hundreds. Deep learning is a type of machine learning in which a model learns to perform classification tasks directly from images. The network increases the complexity and detail of what it is learning from layer to layer. Deep learning requires very large data sets and computationally intensive. However, such approach allows us to learn features and classifiers automatically. The accuracy is unlimited in this case. The network learns directly from the data. We have no influence over what features are being learned [8].

At present, a large number of models of deep neural networks are proposed. A convolutional neural network (CNN) is one of the most popular algorithms for deep learning with images. The first successful applications of Convolutional Networks were developed by Yann LeCun. Its named LeNet [9], Fig. 5.

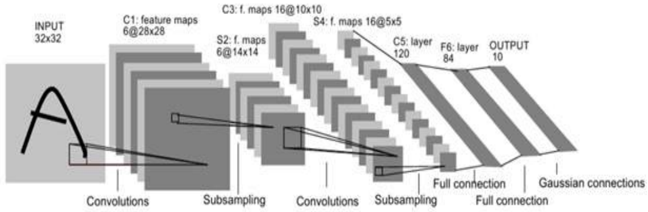

Figure 5. Architecture of LeNet-5, a Convolution Neural Network

The AlexNet was submitted to the ImageNet ILSVRC challenge [10] and significantly outperformed the second runner-up [11], Fig. 6.

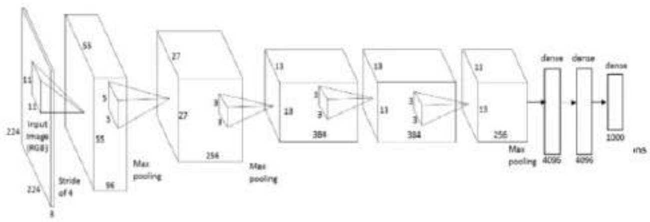

Figure 6. Architecture of AlexNet, a Convolution Neural Network

The ILSVRC 2014 winner was a Convolutional Network from Szegedy et al. from Google (GoogLeNet) [12]. This is not a complete list of models.

The goal of R-CNN (Regional CNN) is to take in an image, and identify where the main objects (via a bounding box) in the image. The first generates category-independent region proposals. These proposals define the set of candidate detections available to detector. The second module is a large convolutional neural network that extracts a fixed-length feature vector from each region. The third module is a set of class-specific linear SVMs [13].

Object detection system is presented in Fig. 7.

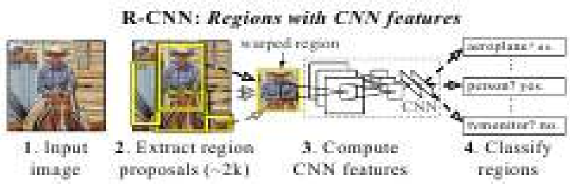

Figure 7. Object detection system [13]

IV. EXPERIMENTAL EVALUATIONS

A few publicly available traffic sign data sets exist:

- German TSR Benchmark (GTSRB) [14,15];
- KUL Belgium Traffic Signs Data set (KUL Data set) [16];
- Swedish Traffic Signs Data set (STS Data set) [6];
- RUG Traffic Sign Image Database (RUG Data set) [17];
- Stereopolis Database [18].

The evaluation of traffic sign detection is based on Swedish Traffic Signs Data set [6, 19]. A dataset has been created by recording sequences from over 350 km of Swedish highways and city roads. A 1.3 mega-pixel color camera at the resolution of 1280×960, a Point-Grey Chameleon, was placed inside a car on the dashboard looking out of the front window. The camera was pointing slightly to the right, in order to cover as many relevant signs as possible. The lens had a focal length of 6.5 mm, resulting in approximately 41 degrees field of view. Typical speed signs on motorways are about 90 cm wide, which corresponds to a size of about 50 pixel if they are to be detected at a distance of about 30 m.

In total, in over 20 000 frames have been recorded of which every fifth frame has then been manually labeled. The label for each sign contains sign type (pedestrian crossing, designated lane right, no standing or parking, priority road, give way, 50 kph, or 30 kph), visibility status (occluded, blurred, or visible) and road status (whether the signs is on the road being traveled or on a side road), see figure 1 [6, 19].

Dataset consists of two subsets. We divided subset1 into training and validation sets consisting of 1970 (20%) images and 7903 (80%), respectively.

In Table II, the results of the implementation of the R-CNN detector for detecting traffic signs for speed limits are presented.

Table I
THE RESULTS OF THE IMPLEMENTATION OF THE R-CNN DETECTOR

Traffic sign	Precision	Recall
Speed limits	$0.833\pm$ 0.01	0.908

In the future, it is planned to conduct experimental evaluations for all sign classes from the used database and expand the test database of images.

In the process of working on a specific task of building intelligent computer vision systems have been obtained acceptable results. However, identified problems and constraints of all tasks in computer vision. Here they are:

- the lack of a unified algorithmic approaches, a large number of individual solutions of a problem;
- testing and analysis of test results on narrow bases;
- the success of the solution depends on the researcher's experience that solves it.

Summarise the foregoing and formulate the problems arising in the process of building computer vision systems in particular, and in the construction of intelligent systems.

V. COMPUTER VISION SYSTEMS: PROBLEMS AND GOALS

Let us generalize what was said above to the general case of constructing an intelligent system of computer vision. We have successfully solved a particular narrow problem.

Until now, research teams have only implemented a method that they believe has potential or perhaps tested a few solutions. This statement is true both for the task of traffic sign detection and for computer vision systems in whole. Without a way to compare performance with other systems, it is not clear which approaches work best. There are a large number of disparate databases, but there is **no knowledge base**.

At present time computer vision systems can perform many tasks and has many real-life applications in a wide range of areas including optical character recognition, face detection, emotion detection, object recognition, vision-based biometrics, identity verification through Iris code, login with fingerprint or face, 3D modeling, special effects, etc. This is a very active research area, and rapidly changing. There are many examples of current computer vision systems. Many software applications have been developed in the last five years.

At present, there are fenced borders, which do not allow speaking about the complex solution of computer vision task. Computer vision system is a particular case intelligent systems in general. The list of limitations in the construction of intelligent systems is large. Let's name only a few of them:

1) There is **no general approach for choosing a method, methodology or algorithm** for solving any computer vision problem. The **experience and knowledge of computer vision specialists are crucial**.
2) Many methods have been proposed for solving problems of computer vision. More sophisticated approaches consist of several computer vision stages. These methods are based on traditional algorithms. Recent works suggest combining different approaches for increasing performance. More often, an increase in system efficiency is achieved by modifying existing methods/algorithms or a combination thereof. Fundamentally new and breakthrough solutions are rarely offered. There are **no universal methods/algorithm for solving wide class of problems**. It follows that there is **no software tools for solving a wide class of computer vision problems**.
3) **Practical vision systems** need to be compact and low cost, but given the previous comments, they are **large-scale and costly**.

The difficulty of choosing an approach for building computer vision system is shown in Table III.

Today there are already a lot of modern technologies of designing intelligent systems. But they can solve not all problems mentioned above. The following statements, formulated for intelligent systems, are also valid for computer vision systems. These are problems to be solved.

1. **The development of intelligent system theory** in general and the theory of computer vision systems in particular is an extremely important task.

2. It is necessary to have a general (complex, integrated, holistic) **technology of designing intelligent systems** for improving effectiveness of designing intelligent systems. Compatibility of such design solutions means compatibility of different kinds of intelligent system components which, in general, can be the products of developing by different and independent developer teams. Intelligent systems based on such technology should be flexible, easy modified, reconfigurable.

3. Development of **different kinds of intelligent system**

Table II
THE DIFFICULTY OF CHOOSING AN APPROACH FOR BUILDING COMPUTER VISION SYSTEM

Component/stage of computer vision system	Technical/algorithmic solution
Visual sensors	**Sensors** Surveillance camera Satellite UAV **Registration method** Monocular Stereo Infra-red Thermal **Video or Image** Video **static or moving** **etc.**
Information extraction	Color features Shape features Temporal features Texture features, **etc.**
Decision-making	Artificial neural networks Support Vector Machine Fuzzy logic Decision trees Bayes classifier Wavelets Hidden Markov Models, **etc.**

components which, in general, can be the products of developing by different and independent developer teams [20].

VI. CONCLUDING REMARKS

In this paper, R-CNN detector is studied as alternative to the typical algorithms of traffic sign detection. Deep learning have improved object detection accuracy. Public traffic signs database was used for experiments.

But this is a special case of solving a narrow problem. Despite the existing advances the problems that exist and which need to be solved in the future in the field of computer vision system design were formulated.

REFERENCES

[1] F. Lu, S. Lee and R. Satzoda, "Embedded Computing Framework for Vision-Based Real-Time Surround Threat Analysis and Driver Assistance," IEEE CVPR - Embedded Vision Workshop, (Best Demo Award), 2016.

[2] A. MØgelmose, M. Trivedi, and T. Moeslund, "Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 4, 2012.

[3] H. Fleyeh, M. Dougherty, D. Aenugula, S. Baddam, "Invariant Road Sign Recognition with Fuzzy ARTMAP and Zernike Moments," Intelligent Vehicles Symposium, 2007 IEEE. pp. 31–36, 2007.

[4] Y.Y. Nguwi, A.Z. Kouzani, "Detection and classification of road signs in natural environments," Neural Comput. Appl. 17, 265–289, April 2008.

[5] Q. Zhang, S.i. Kamata, "Automatic road sign detection method based on Color Barycenters Hexagon model," ICPR 2008.

[6] F. Larsson and M. Felsberg, "Using Fourier descriptors and spatial models for traffic sign recognition," Proc. Image Anal., 2011, pp. 238–249.

[7] Yuan Li, Andreas Møgelmose, and Mohan M. Trivedi, "Pushing the "Speed Limit": High-Accuracy U.S. Traffic Sign Recognition with Convolutional Neural Networks," IEEE Transactions on Intelligent Vehicles, 2016.

[8] MathWorks website [Online]. Available: https://www.mathworks.com

[9] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, 86(11):2278–2324, November 1998.

[10] Large Scale Visual Recognition Challenge (ILSVRC) website [Online]. Available: http://www.image-net.org/challenges/LSVRC/

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in neural information processing systems, 2012.

[12] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, "Going deeper with convolutions". [Online]. Available: https://arxiv.org/pdf/1409.4842.pdf

[13] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation Tech report (v5)". [Online]. Available: https://arxiv.org/pdf/1311.2524.pdf.

[14] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The german traffic sign recognition benchmark: A multi-class classification competition," Proc. IJCNN, 2011, pp. 1453–1460. [Online]. Available: http:// benchmark.ini.rub.de/?section=gtsrb

[15] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," Neural Netw., vol. 32, pp. 323–332, Aug. 2012. [Online]. Available: http:// www.sciencedirect.com/science/article/pii/S0893608012000457

[16] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," Machine Vision and Applications. New York: Springer-Verlag, Dec. 2011, pp. 1–15. [Online]. Available: http://dx.doi.org/10.1007/s00138-011-0391-3.

[17] C. Grigorescu, and N. Petkov, "Distance sets for shape filters and shape recognition," IEEE Trans. Image Process., vol. 12, no. 10, pp. 1274–1286, Oct. 2003.

[18] R. Belaroussi, P. Foucher, J. Tarel, B. Soheilian, P. Charbonnier, and N. Paparoditis, "Road sign detection in images: A case study," Proc. ICPR, Istanbul, Turkey, 2010, pp. 484–488.

[19] Computer Vision Laboratory (CVL) website [Online]. Available: http://www.cvl.isy.liu.se/research/datasets/traffic-signs-dataset/

[20] V. V. Golenkov, "Ontology-based Design of Intelligent Systems" Proc. of the Intenational Conference "Open Semantic Technologies for Intelligent Systems (OSTIS-2017)", Minsk, 16 - 18 February 2017, BSUIR – pp. 37–56, 2017.

ДЕТЕКЦИЯ ДОРОЖНЫХ ЗНАКОВ И ПРОБЛЕМЫ В ОБЛАСТИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Лукашевич М. БГУИР, Минск, Беларусь

Детекция объектов является типовой задачей компьютерного зрения. В данной статье представлены некоторые результаты реализации системы распознавания дорожных знаков. Для детекции использован R-CNN детектор. Ключевое внимание было уделено детекции дорожных знаков ограничения скорости. R-CNN детектор, основанный на глубинном обучении, - это простая и подходящая модель для распознавания дорожных знаков. Этот подход комбинирует низкоуровневые признаки изображения с высокоуровевым контекстом, включающим детекцию объектов и классификатор. Несмотря на существующие подвижки в компьютерном зрении в статье рассмотрены существующие и требующие решения проблемы построения систем компьютерного зрения.