Obstacles and Traffic Signs Tracking System

Anatolii Amarii [1], Stepan Melnychuk [2] and Yuliya Tanasyuk [3]

[1,2,3] *Yuriy Fedkovych Chernivtsi National University, Kotsyubynsky 2, Chernivtsi, 58012, Ukraine*

Abstract

The analysis, development, software implementation and testing of the methodologies for tracking obstacles and road signs have been performed. The created system utilizes artificial neural network of DeepLab for semantic segmentation of the car camera images to identify obstacles and to select traffic signs segments based on MobileNetV2. The TrafficSignNet artificial neural network is subsequently used for traffic signs classification. The software is implemented in the Python programming language using the Tensorflow machine learning platform and the OpenCV, Scipy and Skimage computer vision libraries.

Keywords

Artificial neural networks, semantic segmentation, classification, computer vision

1. Introduction

Nowadays there are many different systems of human assistance in different areas. More and more often the ability to recognize images becomes the requirement for such systems. The problem of image recognition is to identify certain patterns in the picture and relate them to predefined classes.

The driver behind the wheel needs to monitor not only the road conditions, but also the indications of the car sensors, such as current speed, engine RPM, position on the GPS map. Although modern cars are designed so that all the necessary information is available in the driver's field of vision, even occasional distraction from the road to a device can lead to unpredictable consequences.

To solve this problem, road tracking assistant systems are developed. Their operation is primarily based on the algorithms and methods of the road situation analysis with the use of computer vision. The capabilities of such systems include the detection of various obstacles and road signs in the path of the vehicle.

Similar obstacle and road tracking systems were produced at the following companies: Continental (in collaboration with DigiLens Inc.) [1] and WayRay [2]. Both companies have implemented full-fledged hardware and software solutions with the use of augmented reality.

The aim of the given research is to develop a methodology for determining obstacles and road signs in the direction of the car movement with the designation of the entities detected from the video stream in real time on a computer screen.

So, it was decided to explore this area and develop a methodology for analysis of physical objects located within the car route with the use of edge computing. And it will assist in further informing of a vehicle driver and facilitate decision-making.

2. System development methodology

To develop an obstacle and road sign tracking system the classification and semantic segmentation by artificial neural network was used. Artificial neural networks are commonly applied for image processing and show high values both in accuracy and computing speed.

The task of image classification is to determine whether its content belongs to a certain class. In

ISIT 2021: II International Scientific and Practical Conference «Intellectual Systems and Information Technologies», September 13–19, 2021, Odesa, Ukraine
EMAIL: amarii.anatolii@chnu.edu.ua (A. 1); s.melnychuk@chnu.edu.ua (A. 2); y.tanasyuk@chnu.edu.ua (A. 3)
ORCID: 0000-0001-8650-0521 (A. 3)

contrast, semantic segmentation is designed for labelling each pixel of an image correspondingly. Therefore, instead of belonging to one certain class, an image can be related to several categories. As it is shown in Figure 1, classification would determine that there is a cat in the picture. While the segmentation would identify the same image not only as a cat, but also the sky, trees and grass.

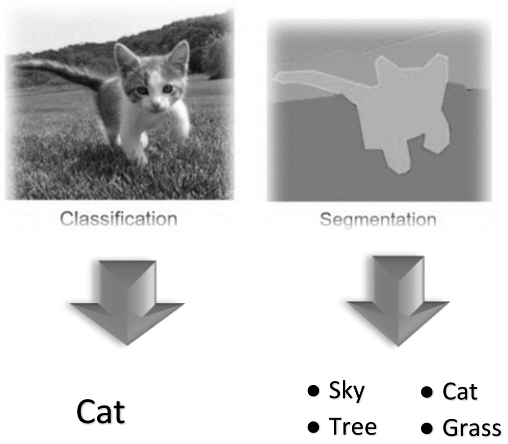


Figure 1: Difference between image classification (left side of picture, where the image of cat is recognized as cat) and image segmentation (right side of picture, where the image of cat is split up on segments of sky, trees, actual cat and grass) [3]

It the developed system neural network semantic segmentation is used for identifying different traffic objects such as other vehicles, people, buildings, trees etc. Also, the created software utilizes capabilities of neural network to detect traffic signs, recognized by means of classification.

2.1. Model of neural network for classification

The analysis of neural network models was performed among those presented on the official GitHub repository of the open machine learning platform TensorFlow [4]. Most of the models considered either require high-power computing systems (such as ResNet and EfficientNet) or were developed for a specific purpose (MARCO). Therefore, it was decided to use a third-party model: a specially created network TraficSignNet for road sign recognition [5]. This type of a network takes advantage of a data set ready-made for training and its simple structure that does not require any complex calculations. The architecture of the deployed neural network is shown in Figure 2.

An image (Input image) with a size of 32x32 pixels with 3 color channels is fed to network`s input. The first convolution layer uses eight 5x5 filters with ReLU activation function and 2x2 aggregation at the maximum value.

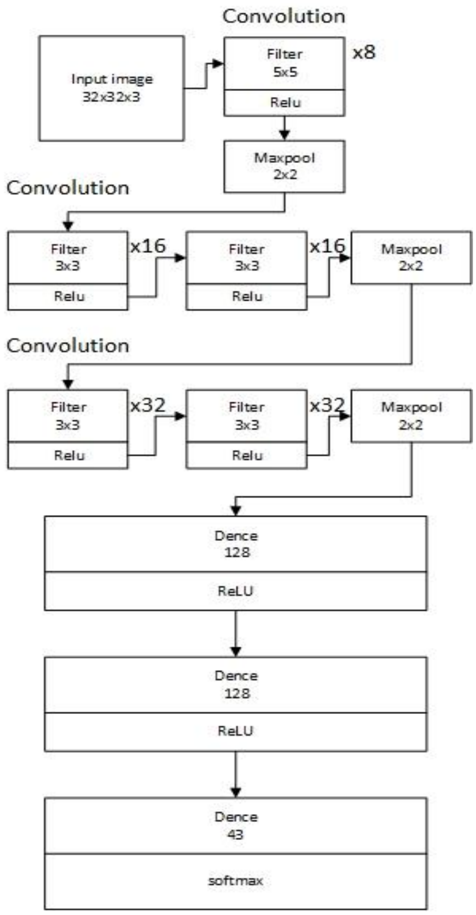


Figure 2: Architecture of TraficSignNet

Each following layer may differ in the number of filters and their dimensions. So, in the next two layers, 16 3x3 filters are used, with the following number of filters increased up to 32. The subsequent three layers are fully connected, where the last one contains 43 neurons, each corresponding to the number of road sign classes in the training data set.

2.2. Model of neural network for semantic segmentation

For image semantic segmentation it was decided to use the Deeplab model [6], which is an example of the "encoder-decoder" architecture.

The encoder is a pre-trained classification network. The MobileNetV2 model was chosen for the encoder network, the architecture of which can be seen in Figure 3.

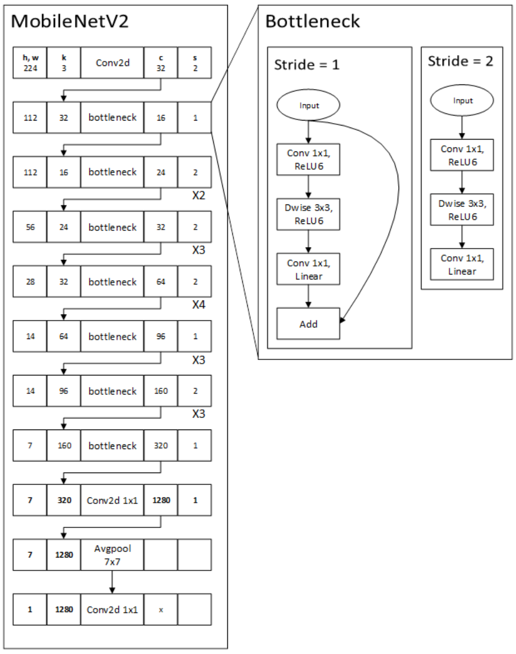


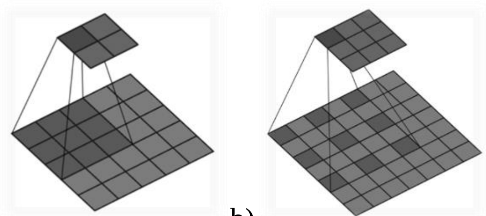
Figure 3: Architecture of MobileNetV2

The MobileNetV2 architecture contains an initial fully convoluted layer with 32 filters, followed by 19 residual bottleneck layers. The ReLU6 activation function is also used to provide nonlinearity due to its reliability when used with low-precision calculations. In addition, we always use 3 × 3 kernel size as a standard for modern networks, and we use screening and batch normalization during training.

In Figure 3 the blocks corresponding to the layers of the neural network contain the following notations: the dimension of the input data (h, w, k), the type of layer (conv2d - convolutional, avgpool - aggregation by the average value), the output number of channels (c), which determines the parameter k of the next layer, and the offset (stride - s), which determines the parameters h and w of the next layer and the structure of the "bottleneck". Below the layers there is the number of repetitions of layers with identical parameters.

DeepLab applies some modifications to this model, changing the ordinary convolution (Fig. 4a) to an atros convolution via kernel dilation rate addition (Fig. 4b) to obtain the characteristics calculated by deep convolutional neural networks with arbitrary resolution. It reduces the calculation time without degrading the accuracy.

The task of the decoding network is to semantically project the discriminant features (lower resolution) learned by the encoder network onto the pixel space (higher resolution) to obtain a dense classification.



a) b)

Figure 4: Types of convolutions: a) ordinary (without dilation rate); b) atrous (with dilation rate) [7]

3. Algorithm traffic signs and obstacles recognition

The algorithm of traffic signs and obstacles recognition implemented in the developed software system is shown in Figure 5. The description of the algorithm is as follows.

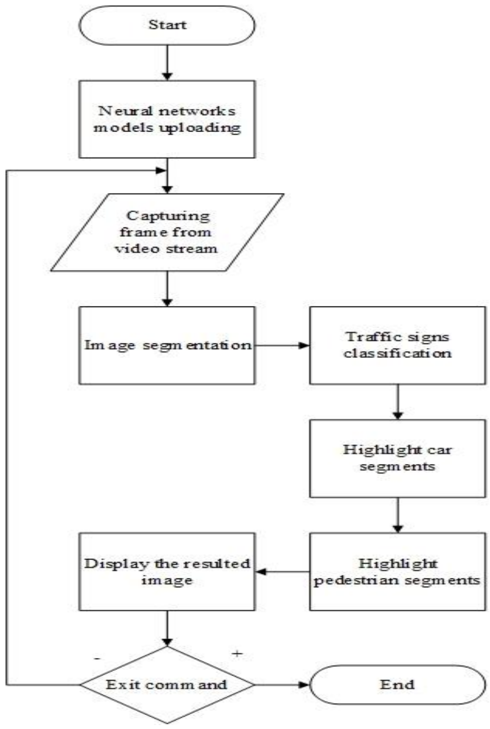


Figure 5: Block diagram of the algorithm

The video camera captures images along the car route (Fig. 6). The image is pre-processed and fed to input DeepLab segmentation model, which returns a segmentation map (Fig. 7).

The resulting segmentation map is divided into segments. Each set of segments is passed for processing to the corresponding module. When the module of road signs classification receives a sample (Fig. 8) it breaks it into separate segments omitting too small objects. After that, each of the remaining segments is further processed and applied as an input to the classification network, resulting in the road sign class definition and its corresponding designation in the frame (Fig. 9).

Figure 6: Image from car camera

Figure 7: Segmentation map

Figure 8: Segments of road signs

Figure 9: The road sign class definition (speed limit)

On the segmentation map in the Modules for selecting a vehicle and a pedestrian all segments related to these objects are highlighted (Fig. 10).

Figure 10: Segments of pedestrians

Then the segments are split up additionally and their spatial characteristics are found. When a ratio of a segment size to the original image size is greater than a value, defined by the spatial characteristics of the segment, outline in the shape of ellipse (Fig. 11) is superimposed on the original image (Fig. 12), and its brightness depends on the aspect ratio.

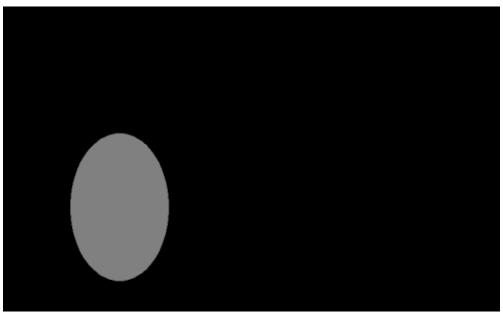


Figure 11: Ellipse, which highlights pedestrians

Figure 12: Resulting image with highlighted pedestrians and detected traffic sign

4. Testing results

The developed method of tracking obstacles and road signs was tested on personal computer equipped with CPU Intel Core i5-2400 and 8GB RAM memory by processing the car's video recordings. The test results have shown that the developed system provides rather small computing time of 0,5 sec, which with an average car speed in the city of 30 km/h is enough to understand the general road conditions and even make decisions. Identification of real pedestrians and cars in the image, distinguishing them from other objects, is performed quite accurately. Although, in the cases when several traffic signs are placed too close to each other, a separate sign can't be clearly distinguished, the system successfully highlights the found segment. However, several shortcomings have also been revealed. Namely, due to the small depth of the artificial neural network for semantic segmentation, extraneous noise objects that do not belong to the specified classes are often distinguished. Moreover, the data set for training an artificial neural network for the classification of road signs contains a fairly limited number of classes (43 entities). In comparison the number of classes in the Ukrainian traffic rules counts 201 entities, excluding plates.

5. Conclusions

The given research considers the application of the means and methods of artificial neural networks for semantic segmentation and image classification with the intention to identify obstacles and perform road signs recognition.

For this purpose, two neural networks have been trained. One of them provides semantic segmentation of images, enabling one to define several entities of different classes as well as their location in a given image. The second neural network is used to recognize road signs.

The test results proved the developed method to be sufficiently effective in identification of physical objects and single road signs located within the car route. Object recognition time is less than 0,5 sec, which implies the use of the proposed method for obstacles detection both in real time and with the video of car recordings. Taking into account the achieved results of testing and utilization, the developed software can be further combined with the facilities of edge computing to provide the driver with notification and decision-making system.

6. References

[1] Continental Group, Augmented-Reality HUD. URL: https://www.continental-automotive.com/en-gl/Passenger-Cars/Information-Management/Head-Up-Displays/Augmented-Reality-HUD-(1)

[2] WayRay. URL: https://wayray.com/

[3] Jasmin Kurtanović, Deep Learning – Semantic Segmentation. URL: https://serengetitech.com/tech/deep-learning-semantic-segmentation/

[4] TensorFlow Model Garden. URL: https://github.com/tensorflow/models

[5] Adrian Rosebrock, Traffic Sign Classification with Keras and Deep Learning. URL: https://www.pyimagesearch.com/2019/11/04/traffic-sign-classification-with-keras-and-deep-learning/

[6] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IV, volume 40 of IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, pp. 834-848.

[7] Paul-Louis Pröve, An Introduction to different Types of Convolutions in Deep Learning. URL: https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d