|  |  |
| --- | --- |
|  |  |

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

Villamosmérnöki és Informatikai Kar

Computer Vision Systems Project Summary

Kürti Ádám (Neptun kód)

Pucsok Sándor PP22N9

2023

**Content**

[**1.** **Project Description** 3](#_Toc136267272)

[**2.** **Car detection algorithm** 4](#_Toc136267273)

[**2.1** **YOLO Algorithm** 4](#_Toc136267274)

[**2.2** **Car detection** 5](#_Toc136267275)

[**3.** **License plate detection and read** 6](#_Toc136267276)

[**4.** **Estimate Vehicle Speed** 8](#_Toc136267277)

[**5.** **Test** 9](#_Toc136267278)

[**6.** **Summary** 10](#_Toc136267279)

[**Literature** 11](#_Toc136267280)

1. **Project Description**

**This project combines research and development in the field of computer vision with algorithms for detecting cars, their license plate and estimate a velocity. Computer vision is a technology that enables computers to perceive, interpret, and understand images or videos.**

**The aim of this project is to develop an algorithm that can efficiently recognize and detect cars, license plates and approximate velocity in video recordings. The algorithm leverages the power of deep learning for object detection and assigning corresponding categories.**

**The foundation of car detection lies in a vast dataset used during the training phase of machine learning algorithms. The algorithm is first trained using images containing a variety of cars and their corresponding labels. Subsequently, it becomes capable of recognizing learned patterns and features in new images, even those not present in the training data.**

**In this summary, we will first present the car detection method, discuss the solutions used, then the license plate recognition algorithm, and finally the algorithm for approximate speed estimation.**



**1**. figure Ide lehetne tenni a működő algoritmusból egy képet

1. **Car detection algorithm**

In this chapter we introduce the algorithm what we use for car detection.

First we tried a solution to train our own neural network, for which we requested the training data from an XY Github user. After training, the network did not work properly, so we had to resort to another solution. The next idea was to use the YOLO algorithm.

## **YOLO Algorithm**

You Only Look Once (YOLO) is one of the most popular model architectures and object detection algorithms. It uses one of the best neural network architectures to produce high accuracy and overall processing speed, which is the main reason for its popularity.

YOLO algorithm aims to predict a class of an object and the bounding box that defines the object location on the input image. It recognizes each bounding box using four numbers:

* Center of the bounding box ((bx, by)
* Width of the box (bw)
* Height of the box (bh)

In addition to that, YOLO predicts the corresponding number c for the predicted class as well as the probability of the prediction (Pc).

Let’s say that we have an image with 2 vehicles, one car and one truck. The first step that YOLO does is dividing the image into a grid.

With the existence of a grid, it’s possible to detect one object per grid cell instead of one object per image. For each grid cell, we can encode a vector that will describe the cell. For instance, the first cell from the top-left doesn’t have any object, and we describe it as:

where (Pc) is the probability of the object class, Bx and By are coordinates of the center of the bounding box, relative to the cell, Bx and By are width and height of the bounding box relative to the whole image, C1 and C2 are 0 or 1 depending on which class represents the bounding box (C1 for car and C2 for truck). Vector (C1, 1) consists of symbols ? because if the first component (Pc) is equal to zero, then the rest of the components can have random numbers are they are not taken into consideration.

Next, if we take the cell that contains the center of the blue bounding box with the cat, we’ll have a vector.

Following this procedure, if we define one vector for each grid cell, the whole image is represented with nine vectors with size 7 or 3x3x7 tensor. This means that in our data set, each image sample is labeled with one 3x3x7 tensor. Using that data set, we are able to create a training and test set and train the convolutional network, which is exactly how YOLO works [1].

* 1. **Car detection**

As mentioned in the previous chapter, the detection was performed using the YOLO algorithm, which was able to achieve accurate detection, and is therefore used in the following.



2. figure YOLO működés közben autókon

As you can see in the picture, the detection is accurate, usually between 85% and 97%, which is really good. As you can see in the picture, the detection is accurate, usually between 85% and 97%, which is really good.

As the detection worked well after several tests, we could move on to the next task, which was to read the number plates.

1. **License plate detection and read**

In this chapter we introduce the algorithm what we use for license plate detection and reading.

As for the car detection, we used the YOLO algorithm. First, we trained the algorithm with different license plates, and after teaching it, we tested how it works in practice. In the next picture. As you can see in the following picture the algorithm works great and detects the license plate numbers accurately.



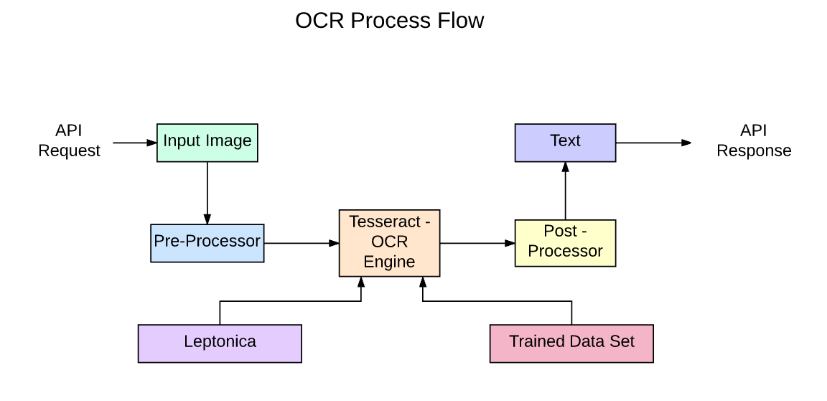
3. figure Kép a rendszám detektálásról

One of the main solutions to the algorithm is to take a picture of the object when you are most likely to see a car. At this point, the car is in the highest focus, so the accuracy of the reading is increased.

* 1. **Reading of license plates**

To read the license plates we use the Tesseract OCR engine. Tesseract 4 adds a new neural net (LSTM) based OCR engine which is focused on line recognition, but also still supports the legacy Tesseract OCR engine of Tesseract 3 which works by recognizing character patterns. Compatibility with Tesseract 3 is enabled by using the Legacy OCR Engine mode (--oem 0). It also needs traineddata files which support the legacy engine, for example those from the tessdata repository.

* Tesseract has Unicode (UTF-8) support, and can recognize more than 100 languages "out of the box".
* Tesseract supports various image formats including PNG, JPEG and TIFF.
* Tesseract supports various output formats: plain text, hOCR (HTML), PDF, invisible-text-only PDF, TSV and ALTO (the last one - since version 4.1.0) [2].



4. figure OCR Process Flow [3]

To see the accuracy of the algorithm, we tested it on several of our own videos, the following table shows the data from one such test:

|  |  |
| --- | --- |
| **Original License Plates** | **Read from Algorithm** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

1. **Estimate Vehicle Speed**
2. **Test**

When the algorithm was fully ready, we shot several test videos in the city with an Olympus OMD E-M10 camera. The most suitable location was Petőfi Bridge, Buda side, as we could film cars driving along the quay from the bridge from above.

However, the camera had a property that only objects closer to it were sharp enough, so a mask was added to the video so that detection, reading and speed measurement only occurred in this area, speeding up the algorithm and optimizing it.



5. ábra kép a maszkolásról

We made several tests and processed them, and based on them we can say that the algorithm meets our current learning goal and 90% of what we had in mind at the beginning of the project.

1. **Summary**

This project, in which we developed an algorithm for car detection, license plate recognition, and speed estimation, has been an incredible experience for us. Over the past period, we have put in intensive work to understand the principles of computer vision, the latest methods in object detection and character recognition, as well as techniques for speed estimation.

Working on this project not only enhanced our theoretical knowledge but also provided us with practical experience. We encountered challenges and difficulties along the way, which we successfully overcame. Preparing the data, debugging errors, and fine-tuning the algorithm were all tasks in which our team collaborated effectively.

In conclusion, this project has been a rewarding journey that has allowed us to delve into the fascinating field of computer vision and machine learning. We have gained valuable knowledge and skills that will undoubtedly be beneficial in our future endeavors.

# **Literature**

[1] What is YOLO: <https://www.baeldung.com/cs/yolo-algorithm>

[2] Tesseract OCR: <https://github.com/tesseract-ocr/tesseract>

[3] Picture of OCR Process Flow: <https://nanonets.com/blog/ocr-with-tesseract/>