TRACKING THE FLOW OF MOTOR VEHICLES ON THE ROADS WITH YOLOV5 AND DEEPSORT ALGORITHMS

Kuchkarov T. A., Hamzayev J. F., Allamuratova Z. J.
Department of Computer systems, Department of information computer technologies and programming
TUIT named after Muhammad al-Khwarizmi Tashkent, joint faculty TUIT-BSUIR
Tashkent, Uzbekistan
E-mail: {timanet4u, hamzayevjf, zamira.lars}@gmail.com

This article describes one of the most important steps in smart transportation systems, which is to count the flow of moving vehicles, in particular, using the YOLOv5 and DeepSORT algorithms to perform this process. It explains how much accuracy increases when these algorithms are used in situations. Experiments using the algorithms used showed an accuracy of 94. One of the most important steps in the development of intelligent transport systems today is the classification of vehicles moving on the roads and counting their flow. The focus of this study is on counting the flow of traffic on the roads. detects that the vehicle has passed. In addition, this study focuses on the process and mathematical aspects of the YOLOv5 and DeepSORT algorithms, with examples. In particular, attention is paid to the Mahalanobis distance and the Kalman filter, which are the main components of the DeepSORT algorithm, and their role in the DeepSORT algorithm is revealed. Experiments based on video images taken from a CCTV device using these algorithms are presented in the article. The main part of the results of the experiments is presented in the article as an example.

Introduction

Nowadays, in the period of rapid development, the transport system has become an important part of human life. Just one example of this is the following statistic: by the end of the first quarter of 2022, it is known that there will be approximately 1.45 billion cars in the world, of which about 1.1 billion will be passenger cars. This means that there is a car for every 7.18 people on our planet, it should be noted that at the time of writing this article, the population of the planet earth exceeds 7.9 billion [1, 9].

Based on the information mentioned above, it can be said that this will definitely cause traffic problems in big cities. In this research work, the analysis of the algorithms used in the program that performs one of the main roles for the system helping to solve this traffic problem, that is, counting the flow of motor vehicles on the road and identifying the moving motor vehicles, was carried out. In particular, YOLOv5 and DeepSORT algorithms are used and the working process of these algorithms is explained.

I. Application of YOLOv5 and DeepSORT algorithms in the calculation of traffic flow

YOLOv5 algorithm. Nowadays, object detection is one of the most important aspects of computer vision. Object detection is a computer vision technology that localizes and identifies objects in an image.

Due to the versatility of object detection, object detection has emerged as the most widely used computer vision technology in the last few years [8]. Let's talk briefly about how object detection works in general. In the field of computer vision, object detection is often referred to as image detection, and these concepts are generally considered synonymous terms. Object detection is not similar to other common computer vision technologies, such as classification, keypoint detection, or semantic segmentation, that is, image segmentation using masks. This is one of the most comprehensive areas of computer vision.

This is where object detection models come into play and achieve object detection by predicting the X1, X2, Y1, Y2 coordinates and object class labels depending on the target. Using object detection simply involves inputting an image (or video frame) into the object detection model and outputting a file in the form of a JSON file with estimated coordinates and class labels [8].

The YOLO (You Only Look Once) algorithm is a modern, real-time, very effective object detector, and there are a number of algorithms included in this group of algorithms (YOLOv1, YOLOv2, YOLOv3, YOLOv4, YOLOv5). The algorithm used to obtain the results in this article is the YOLOv5 algorithm.

It can be observed that the YOLOv5 algorithm is based on YOLOv1-YOLOv4. Continuous updating of this algorithm allowed to achieve high results in two official databases of object detection, i.e. Pascal VOC (visual object classes) and Microsoft COCO (common objects in context). in the database [5,6]. The network architecture of YOLOv5 is shown in Figure 1.

There are several other reasons why YOLOv5 was chosen for this research work. First, YOLOv5 introduced a cross stage partial network (CSPNet) into the Darknet and created CSPDarknet as its foundation [12]. Here Darknet is a CNN (Convolutional Neural Network), i.e. a convolutional (sliding) neural network used in the YOLO algorithm. CSPNet solves the problem of repeated gradient data on large-scale networks and integrates gradient transformations into a feature map, thereby reducing model parameters and FLOPS (floating-point operations per second), which not only improves inference speed and accuracy. provides, but also reduces the size of the model. Detection speed and accuracy are very important in the task of traffic flow detection on roads, and the compact size of the model also determines its inference efficiency in resource-poor peripheral devices.

Second, YOLOv5 used a path aggregation network (PANet) to enhance information flow [7]. PANet adopts a new feature pyramid network (FPN) structure with an improved bottom-up path that improves the propagation of low-level features.

Third, the YOLOv5 head, the YOLO layer, generates feature maps of 3 different sizes (18×18, 36×36, 72×72) to achieve multiscale prediction, which allows the model to handle small, medium, and large object sizes [4].

Here:

- Cross Stage Partial Network (CSP) - cross stage partial network;
- Spatial Pyramid Pooling (SPP) - spatial pyramid pooling;
- Convolution Layer (Conv) - convolution (sliding) layer;
- Concatenate Function (Concat) - function of combining.

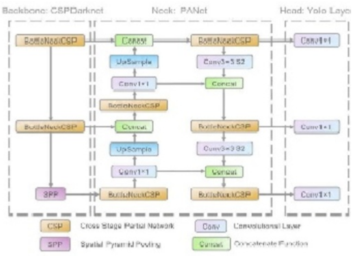

Figure 1 – Yolov5 network architecture

DeepSORT algorithm workflow

Taking into account what Kalman filter and Mahalanobis distance are given above, we will try to explain the working process of DeepSORT algorithm. DeepSORT technology combines these two concepts to transfer information from one frame to another and adds a new metric called appearance. In the first step, the position, size, and class of a single bounding box are determined using object detection. Then, in principle, the Hungarian algorithm (Hungersky algorithm) can be applied to associate previously known objects with object identifiers previously in the frame and observed using Kalman filters. As a result, everything works perfectly as in the original SORT.

II. Conducted experiments and obtained results

Below, using YOLOv5 and DeepSORT algorithms mentioned above, the movement of cars on the roads was determined and exactly how many and what type of vehicles passed through a certain part of the road during the specified time. and good results were recorded. Based on the obtained results, the stages of the algorithm execution are presented as examples in the form of an image clipped from the video:

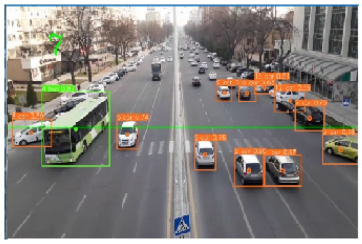

Figure 2 – Samples of the results obtained on the basis of the experiment

Conclusion

This article describes the process of identifying vehicles moving on the road using the YOLOv5 and DeepSORT algorithms, as well as the algorithm for tracking and counting their movements. Experimental results based on video images obtained from a CCTV device installed in a designated area have shown that this algorithm performs very well in real time. In particular, DeepSORT's ability to remember an object by its appearance and track it through these features is a very effective method, as this was not available using the SORT algorithm itself until now. The proposed algorithm can be considered as an effective method for the algorithm because it consumes less resources in the calculation process and there is no need to train the model.

III. References

1. Kuchkarov T. A., Xamzayev J. F., Ochilov T. D. Intelektual transport tizimi ilovalari uchun sun'iy intelekt texnologiyalaridan foydalanish // Berdaq nomidagi QDU Axborotnomasi, no. 2 (51) 2021, 114–120 b.
2. Zhang L.,Li Y. and Nevatia R. Global data association for multi-object tracking using network flows // in CVPR, 2008, pp. 1–8.
3. Berclaz J., Fleuret F., Turetken E. and Fua P. Multi-ple object tracking using k-shortest paths optimization // IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 9, pp. 1806–1819, 2011
4. Yang B. and Nevatia R. An online learned CRF model for multi-target tracking // in CVPR, 2012, pp. 2034–2041.
5. Yang B. and Nevatia R. Multi-target tracking by on-line learning of non-linear motion patterns and robust appearance models // in CVPR, 2012, pp. 1918–1925.
6. Andriyenko A. , Schindler K. and Roth S. Discrete-continuous optimization for multi-target tracking // in CVPR, 2012, pp. 1926–1933.
7. Solawetz J. How to Train YOLOv5 On a Custom Dataset.// https://blog.roboflow.com/author/jacob/
8. Ishmuratov T. Kak rabotaet Object Tracking na YOLO i DeepSort // https://habr.com/ru/post/514450/
9. Lin T. Y., Maire M., Belongie S. Hays J., Perona P., Ramanan D., Dollár P., Zitnick C. L. Microsoft coco: Common objects in context. //In Proceedings of the 13th European Conferens.