4th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Heavy vehicle detection using optical remote sensing images and deep learning

Roya Talebi*[1], Sadra Karimzadeh [1], Gordana Kaplan [2]

[1]University of Tabriz, Department of Remote Sensing and GIS, Tabriz, Iran
[2] Eskisehir Technical University, Institute of Earth and Space Sciences, Eskisehir, Türkiye

Keywords
Remote sensing
Object detection
YOLOv5
CNNs
Heavy vehicle

Abstract

Recently, object detection has experienced significant progress. However, in the field of remote sensing, it is still under investigation for real-time detection of small objects because of the limited resolution and information. The detection of a predefined object, such as heavy vehicle, is vital in some applications. This study aims to improve the detection accuracy of the heavy vehicle. We train the latest model, YOLOv5 on our dataset in this paper. The results show that YOLOv5 can be well applied in the field of heavy vehicle detection.

1. Introduction

Today, the automobile industry of any country is one of the essential indicators for the development of that country. This industry is considered a key industry in terms of extensive relationship with its upstream and downstream chains. It has an important place in industrial development and production prosperity as it plays an effective role in economic growth and development. Thus, reviewing the activities of these industrial poles with the help of remote sensing data and techiqes can significantly reduce the need for human monitoring and save time and money for various organizations. On the other hand, this method can be used to check other collections as well. For example, identifying heavy goods vehicles in rest areas (Kasper-Eulaers et al., 2021).

A large number of remote sensing images have been generated regularly, and due to the rapid development of satellite and imaging technology, the task of object detection has gained significant attention of researchers (Zakria et al., 2022).

In this paper, we used the latest version of the You Only Look Once (YOLO) object detection algorithm (Jocher et al., 2020) to detect vehicles. Our focus was on applying the algorithm, data acquisition, data annotation and Data counting.

YOLOv5 is the fifth generation of YOLO, written in Python programming language (Thuan, 2021). According to various studies YOLOv5 outperforms the rest of the YOLO model in terms of accuracy and speed (Thuan, 2021); (Cengil & Çinar, 2021).

YOLOv5 proposed by Ultralytics LLC is an improved version based on YOLOv4. It is a one-stage detection network regarding accuracy and detection speed (Wu et al., 2020). After learning from the advantages of the previous version and other networks, YOLOv5 changes the characteristics of the previous YOLO target detection algorithm that the detection speed is faster but the accuracy is not high. YOLOv5 has improved detection accuracy and real-time performance, which meets the needs of real-time image detection and has a smaller structure. Therefore, this article uses YOLOv5 as the detection model. Its network model is divided into 4 parts, namely Input, Backbone, Neck and Prediction, and its network structure is shown in Fig. 1 (Tan et al., 2021). Input includes mosaic data enhancement, adaptive anchor frame calculation, and adaptive image scaling. The input terminal of YOLOv5 adopts the same mosaic data enhancement method as YOLOv4. The random clipping, scaling, and distribution are used to splice the images. The four images are spliced, which enriches the detection data set, improves the robustness of the network, reduces the calculation of GPU, and increases the universal applicability of the network; Adaptive anchor frame calculation sets the initial anchor frame for different data sets, outputs the prediction frame based on the initial anchor frame, and then compares it with the real frame. After calculating the gap, it updates the network parameters reversely and iterates the network

* **Corresponding Author**

*(roya.talebi1399@gmail.com) ORCID ID 0000-0002-1553-1678
(sadra.karimzadeh@gmail.com) ORCID ID 0000-0002-5645-0188
(kaplangorde@gmail.com) ORCID ID 0000-0001-7522-9924

Cite this study

Talebi, R., Karimzadeh, S., & Kaplan, G. (2022). Heavy vehicle detection using optical remote sensing images and deep learning. 4th Intercontinental Geoinformation Days (IGD), 250-253, Tabriz, Iran

parameters continuously. The anchor frame parameters are (Gao et al., 2021). Adaptive image scaling is to scale the image to a uniform size, which has been implemented in the data preprocessing stage (Luo et al., 2022).

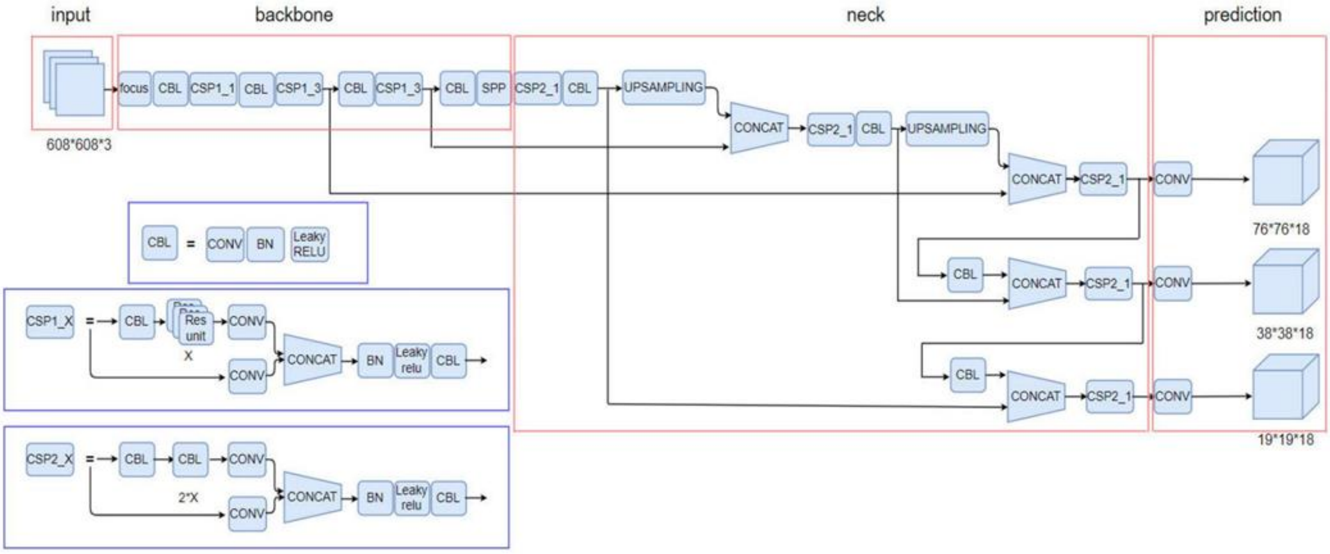

Figure 1. The main modules of YOLOv5 network

2. Materials and Method

2.1. Study area

Tehran is located within latitude of 35°40′18″ and longitude of 51°25′27″ with an altitude of 1,191 m above the mean sea level in a semi-arid region of Iran. Iran Khodro Diesel Company is the study area in this article, which is located in Tehran.

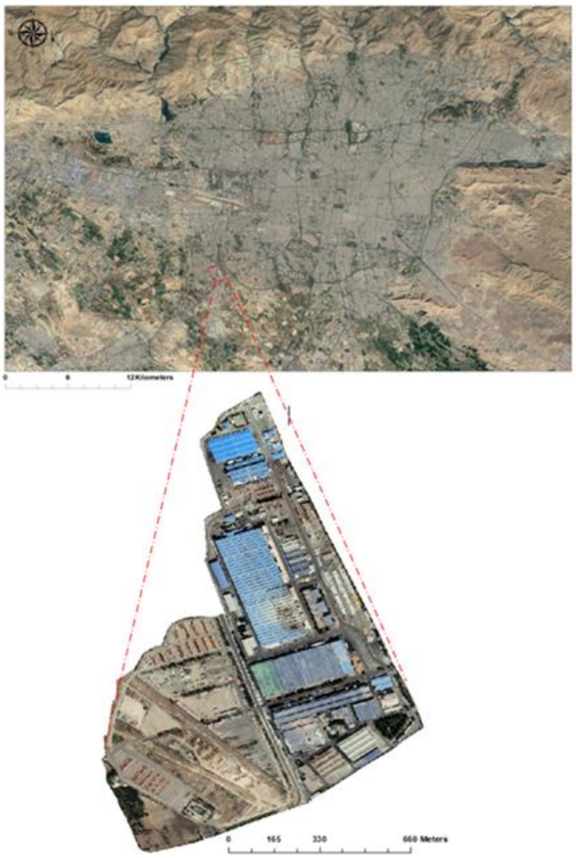

Figure 2. Study Area

2.2. Selection of Algorithm

The decision to use a convolutional neural network (CNN) was made due to their ease of use. There are a number of pre-trained models that can be tuned for a variety of tasks. They are also readily available, computationally inexpensive and show good performance metrics. Object recognition systems from the YOLO family are often used for vehicle recognition tasks, e.g., (Fachrie, 2020); and have been shown to outperform other target recognition algorithms (Benjdira et al., 2019). YOLOv5 has proven to significantly improve the processing time of deeper networks (Jocher et al., 2020).

2.3. Evaluation metric

In order to evaluate performance of the algorithm, results of the model should be compared to the ground truth. In this paper, Precision and recall are used to evaluate the similarity and diversity between detection results and ground truth in test dataset. Also, the overall accuracy and F-Score have been calculated for the accuracy assessment (Kaplan et al., 2021).

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

Where,

TP = true positive
FP = false positive
FN = false negative
TN = True negative.

3. Experiments

3.1. Experimental process

The experimental process can be divided into four steps:

(i) Dataset construction. The remote sensing images (gain from open-source data source such as GoogleEarth are animated using imgLabel to obtain standard YOLO format dataset. Tzutalin's (Tzutalin, 2018) popular annotation tool LabelImg was used to cautiously label the images in this final phase. First, each image is opened in this tool one at a time. Then, a rectangular shape was manually drawn to the boundary of an object in order to specify its exact location in that image by x_center y_center width height. Finally, each object has been given a label, such as 'heavy vehicle'. In LabelImg, annotated values were saved as txt files in YOLOv5 format. Divide labeled dataset into training, validation, and test sets.

(ii) Model construction. Constructing a CNN structure and setting its hyperparameters.

(iii) Model training. Training with training sets and validation sets.

(iv) Model prediction. Testing with the test set, and the result is used for evaluates model.

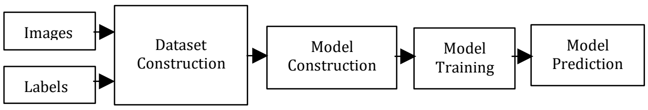

Figure 3. Experimental Process

3.2. Dataset

For this study, 1000 images of the study area were prepared by Google Earth Pro software. Imglabel software is used for manual annotation of heavy vehicles in the images. Fig. 4 shows one demos of image label. 800 images have been used for practice. To test the work, 20% of the images selected for the exercise and another 150 images from the image collection have been selected. Finally, the proposed method is tested with the remaining 50 images.

Figure 4. Demo of image label

3.3. Training

The model was trained using Google Colab, which provides free access to powerful GPUs and requires no configuration. We used a notebook developed by Roboflow.ai (Li et al., 2021) which is based on YOLOv5 (Jocher et al., 2020) and uses pre-trained COCO weights. We added the desired area dataset and adjusted the number of epochs to be trained as well as the stack size to train the upper layers of the model to detect our classes. Training a model for 500 epochs takes about 150 min.

3.4. Experimental Analysis

After training our model, we made predictions for our test set's new and unseen pictures. The examples in Fig. 5 show that the algorithm can detect the heavy vehicle to a higher degree of certainty. However, it has difficulty the Data counting, especially when the data is in the corner of the image.

3.5. Results

The YOLOv5 model performs really well with 230 epochs. After that with the increase of epochs all the losses like classification loss, box loss and objectness loss were increased and the model performance was decreased. YOLOv5 was tested on other images to detect heavy vehicles. To train the model, various image resolutions were used.

The mold dataset was designed to train by using Google Colab, which provides free access to powerful GPUs. We used a notebook developed by Roboflow.ai which is based on YOLOv5 (Roboflow, 2016) and uses pre-trained COCO weights. Suitable number of epochs was chosen to train newly developed mold dataset. To train the model 205 epochs was selected which was taken approximately 40 min.

Figure 5. Results of heavy vehicle detection

According to the results shown in Table 1, the proposed method has identified the heavy vehicle in the study area with high accuracy. A total of 45 heavy vehicles were identified in the first image and 64 heavy vehicles in the second image. The number of TP (heavy vehicles is correctly identified) was 43 for the first image and 59 for the second image. In order to evaluate the results, accuracy assessment was performed using Equations 1 and 2. The values of P (Precision) and R (Recall) for the first image were 1.0 and 0.95, respectively, and for the second image were 1.0 and 0.92, which indicates the efficiency of the proposed method in this field.

Table 1. Performance of heavy vehicle detection

Items	TP	TN	FP	FN	P	R	OA	F-s
a	43	0	0	2	1.0	0.95	0.95	0.97
b	59	0	0	5	1.0	0.92	0.92	0.95

Results and performances of heavy vehicle detection are shown in Fig. 5 and Table 1.

4. Discussion

Although tremendous progress has been made in the field of object detection recently, it remains a difficult task to detect and identify objects accurately and quickly. Yan et al. (2021) named the YOLOv5 as the most powerful object detection algorithm in present times.

We see the greatest potential for improving performance in adjusting the physical data collection and in improving the data annotation. For most applications, changes to the physical data collection cannot be influenced.

However, as this is a pilot project running on only one Automotive company, there is the possibility of changing the physical setup for data collection if more Automotive companies are added.

In order to get closer to the goal of identifying all the heavy vehicles in the image, we need to train the sample with more images.

5. Conclusion

In order to achieve the accurate and real-time intelligent detection of heavy vehicles in Google Earth images, a real-time heavy vehicle detection method based on YOLO v5 was proposed in this study. The method of using YOLO v5 algorithm was proposed to identify heavy vehicles in the study area. According to the results shown in Fig. 5, the proposed method is able to detect heavy vehicles with an accuracy of 78%. We are confident that with a bigger training set and the implementation of the changes suggested in Section 4, the algorithm can be improved even further.

References

Benjdira, B., Khursheed, T., Koubaa, A., Ammar, A., & Ouni, K. (2019). Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3. 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS),

Cengil, E., & Çınar, A. (2021). Poisonous mushroom detection using YOLOV5. Turkish Journal of Science and Technology, 16(1), 119-127.

Fachrie, M. (2020). A Simple Vehicle Counting System Using Deep Learning with YOLOv3 Model. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 4(3), 462-468.

Gao, S.-H., Han, Q., Li, D., Cheng, M.-M., & Peng, P. (2021). Representative batch normalization with feature calibration. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Jocher, G., Stoken, A., & Borovec, J. (2020). NanoCode012, ChristopherSTAN, L. Changyu, Laughing, tkianai, A. Hogan, lorenzomammana, yxNONG, AlexWang1900, L. Diaconu, Marc, wanghaoyang0106, ml5ah, Doug, F. Ingham, Frederik, Guilhen, Hatovix, J. Poznanski, J. Fang, L. Yu, changyu98, M. Wang, N. Gupta, O. Akhtar, PetrDvoracek, and P. Rai,"ultralytics/yolov5: v3, 1.

Kaplan, O., & Kaplan, G. (2021). Response Spectra-Based Post-Earthquake Rapid Structural Damage Estimation Approach Aided with Remote Sensing Data: 2020 Samos Earthquake. Buildings 2022, 12, 14.

Kasper-Eulaers, M., Hahn, N., Berger, S., Sebulonsen, T., Myrland, Ø., & Kummervold, P. E. (2021). Detecting heavy goods vehicles in rest areas in winter conditions using YOLOv5. Algorithms, 14(4).

Li, B., Fu, M., & Li, Q. (2021). Runway Crack Detection Based on YOLOV5. 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT),

Luo, S., Yu, J., Xi, Y., & Liao, X. (2022). Aircraft Target Detection in Remote Sensing Images Based on Improved YOLOv5. IEEE Access.

Roboflow. How to Train YOLOv5 on Custom Objects. 2016. Available online: https://colab.research.google.com/drive/1gDZ2xcT OgR39tGGs-EZ6i3RTs16wmzZQ (accessed on 3 February 2021)

Tan, S., Bie, X., Lu, G., & Tan, X. (2021). Real time detection of personnel mask wearing based on yolov5 network model. Laser magazine, 42(2), 147-150.

Thuan, D. (2021). Evolution of yolo algorithm and yolov5: the state-of-the-art object detection algorithm.

Tzutalin. 'LabelImg'. 2018. Available at: https://github.com/tzutalin/labelImg. (Accessed on November 2020)

Wu, D., Lv, S., Jiang, M., & Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. Computers and Electronics in Agriculture, 178, 105742.

Yan, B., Fan, P., Lei, X., Liu, Z., & Yang, F. (2021). A real-time apple targets detection method for picking robot based on improved YOLOv5. Remote Sensing, 13(9), 1619.

Zakria, Z., Deng, J., Kumar, R., Khokhar, M. S., Cai, J., & Kumar, J. (2022). Multi scale and direction target detecting in remote sensing images via modified YOLO-v4. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.