DEVELOPMENT OF A SYSTEM FOR SEGMENTATION OF NATURAL TYPES OF UNDERLYING SURFACES CHARACTERISTIC FOR THE AUTUMN SIBERIAN URBAN TERRAIN

Ott G.V.
Tomsk Polytechnic University, School of Computer Science & Robotics, student 8E92, e-mail: grisha-ott@gmail.com

Introduction

The task of segmentation is one of the priority tasks of Computer Vision (CV) in robotics. The segmented image can be used for subsequent construction of robot movement routes or for more accurate object detection. Nowadays, convolutional neural networks have proven to be most successful in solving CV problems.

The aim of this work is to develop a system for segmenting underlying surfaces in natural conditions based on a convolutional neural network implemented using the Detectron 2 framework [1].

Neural network architecture

As part of this work, the COCO-InstanceSegmentation mask_rcnn_R_50_FPN_3x model was used. This model is a Mask R-CNN model, which allows it not only to detect objects, but also to determine whether individual pixels belong to these objects.

The model (mask_rcnn_R_50_FPN_3x) was chosen for a number of parameters, such as the minimum required video memory for training and high box AP and mask AP compared to other models presented in Detectron 2 [2].

Preparing a dataset for model training and validation

The dataset was collected in the city of Tomsk in Autumn. In the course of data collection, 6 main classes of underlying surfaces were identified. These classes are shown in Figure 1.

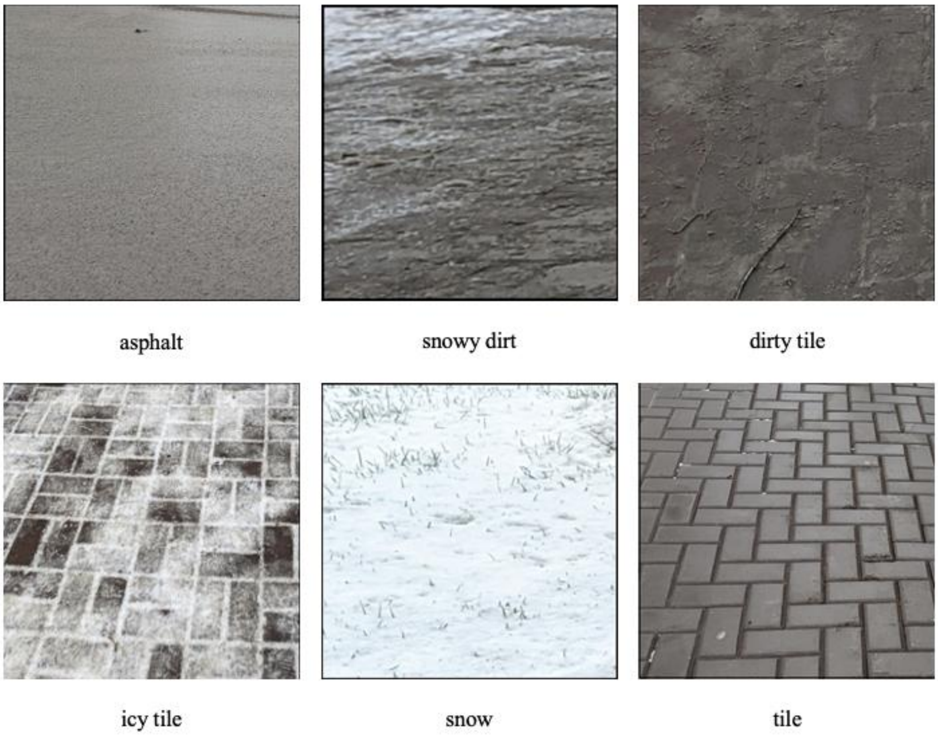

Fig. 1. Selected surface classes

The dataset includes 350 images. All frames were scaled to 16:9, 1920:1080. The training and test sets were separated by 80% and 20%, respectively.

To train the model, you need to prepare the data - select segments in the image and create json files with descriptions of the segments and images. The labelme[3] utility was used for this.

Model training and validation

The model was trained in Google Colab [4]. This site allows working in Jupyter Notebook on Google's cloud servers. The site can allocate up to 15 GB of video memory for operation. At the same time, the platform sets a limit on the time of continuous operation (up to 3 hours).

The model was trained in Google Colab [4]. This site allows you to run Jupyter Notebook on Google's cloud servers. The site can allocate up to 15 GB of video memory for operation. At the same time, the site sets a limit on the time of continuous operation (up to 3 hours).

In connection with this limitation of the operating time, the parameter of the number of iterations of the training cycle (5000) was selected. Due to the limitations of the amount of video memory, a batch size of 8 was chosen, which makes it possible to achieve a fairly smooth change in the gradient on the sample.

During the training, the following metrics were taken, presented in Figures 2-4.

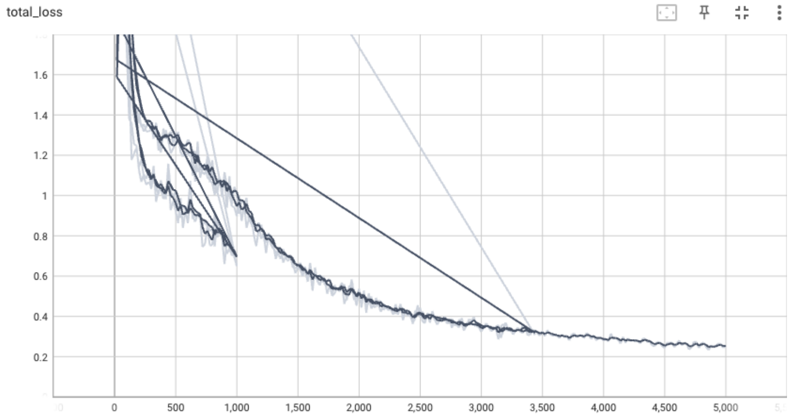

Fig. 2. The value of the total loss during training

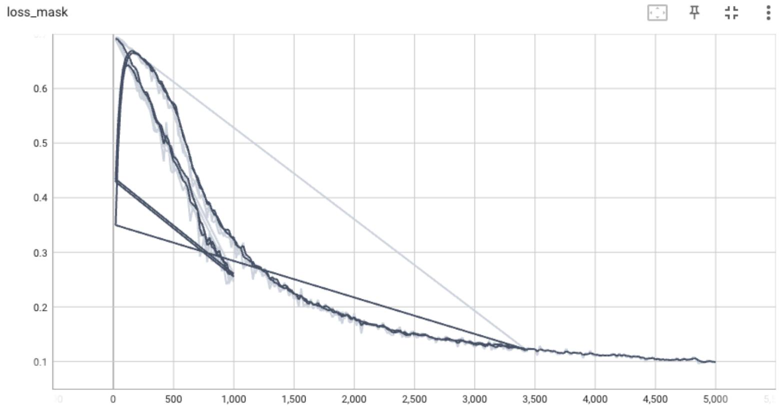

Fig. 3. The value of the loss-function for the class membership mask during training

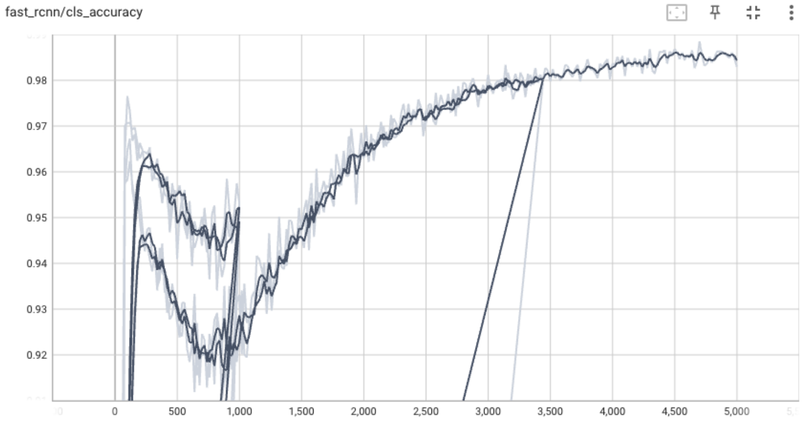

Fig. 4. Change of the class definition accuracy during model training

During the validation of the model on the test dataset, the following metrics were obtained (Table 1).

Table 1

System accuracy for each class

Class	Number of instances	Accuracy
Asphalt	58	86 %
Snow	59	72 %
Tile	59	80 %
Dirty tile	16	72 %
Icy tile	21	70 %
Snowy tile	33	72 %

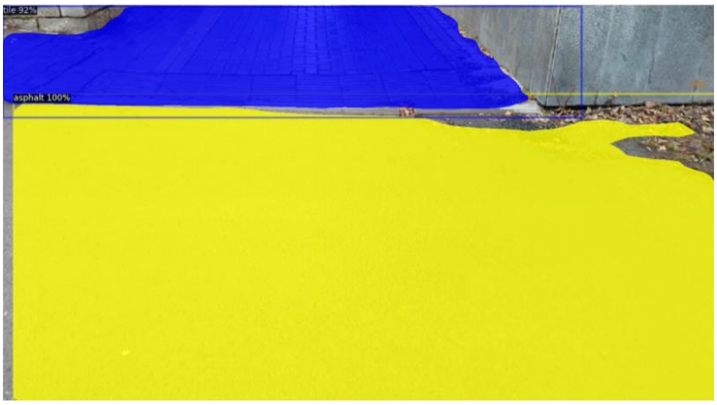

Fig. 5. An example of the operation of the segmentation system

Conclusion

This work demonstrated the possibility of using Mask R-CNN models in segmentation of underlying surfaces in natural conditions.

The resulting model can be improved by preparing a larger dataset and increasing the number of training iterations.

References

1. Detectron2 . – URL: https://github.com/facebookresearch/detectron2 (date of access to the source 09.01.2023).
2. Detectron2 ModelZoo and Baselines. [Электронный ресурс]. – URL: https://github.com/facebookresearch/detectron2/blob/main/MODEL_ZOO.md (date of access to the source 10.01.2023).
3. Labelme. – URL: https://github.com/wkentaro/labelme (date of access to the source 01.02.2023).
4. Google Colab. – URL: https://colab.research.google.com (date of access to the source 05.02.2023).