REVIEW OF MOST COMMON APPROACHES TO OBJECT DETECTION FROM TWO-DIMENSIONAL IMAGES

bachelor Valentyna Yakovyn, Vasyl Stefanyk Precarpathian National University, Ivano-Frankivsk

With the development of robotics, it has become possible to solve a number of tasks arising in household and industrial spheres, particularly in terms of human (user) interaction with the computing system. One important task is the need to ensure satisfactory results in recognizing similar moving objects, including the recognition of letters, barcodes, vehicle numbers, faces, speech, and images [1].

The most common types of recognition tasks used are: identification, which involves distinguishing a certain specific object among similar ones; classification, assigning an object to one class or another; clustering, which involves dividing a given set of objects into classes - groups of objects similar to each other by some criterion. Referring an object to a certain class reflects the most typical classification problem, and when talking about image recognition, this problem is most often implied. Clustering is often called unsupervised classification, as, unlike the previous task, the classes are unknown [2].

As a result, software approaches (libraries, frameworks) that offer ready-made solutions for solving the above-mentioned tasks have undergone intensive development. One such option is You Only Look Once (YOLO) [3], this approach to object detection reformulates it as a single regression problem, directly predicting bounding boxes and class probabilities from full images in a single assessment. Unlike similar object detection systems that adapt classifiers or apply complex pipelines involving multiple stages such as resizing the image, running a convolutional network, and applying non-maximum suppression, the YOLO architecture consists of a single convolutional network that simultaneously predicts several bounding boxes and class probabilities for these boxes. The system divides the input image into a grid and predicts bounding boxes and probabilities for each grid cell, integrating detection and classification into a single unified model, which accelerates the detection process [3].

References: 1. *V. Yakovyn and T. Kriukov*, "Object detection in mobile robotics," 2022 International Conference on Innovative Solutions in Software Engineering (ICISSE), Ivano-Frankivsk, Ukraine, Nov. 29-30, 2022, pp. 310-312, doi: 10.5281/zenodo.8418713 **2.** *M. Kozlenko, V. Sendetskyi, O. Simkiv, N. Savchenko, and A. Bosyi*, "Identity documents recognition and detection using semantic segmentation with convolutional neural network," CEUR Workshop Proceedings, vol. 2923, Kyiv, Ukraine, Jan. 28, 2021, pp. 234-242 **3.** *J. Redmon, S. Divvala, R. Girshick, and A. Farhadi*, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91