УДК 004

VISUAL TARGET DETECTION ALGORITHM YOLO AND ITS APPLICATION AREAS

Guo Qicheng, Li Bohan
Belarusian State University of Informatics and Radioelectronics
e-mail: 2733064287@qq.com

Summary. *YOLO (You Only Look Once) is an object detection algorithm based on deep neural networks, which is used to identify and locate multiple objects in real time in images or videos. The main features of YOLO are fast speed and high accuracy, which can achieve fast object detection in real-time scenarios. YOLO is a fast and accurate object detection algorithm, which is widely used in computer vision fields, including real-time video analysis, autonomous driving, security monitoring, intelligent transportation, defect detection, etc.*

YOLO transforms the object detection task into a regression problem, dividing the image into multiple grids and predicting whether each grid contains an object, as well as the bounding box and category of the object. This means that YOLO only needs a single forward propagation to complete object detection and classification at the same time, thus achieving real-time performance [1].

YOLO consists of network architecture, basic network, feature extraction layer, grid division, bounding box prediction, non-maximum suppression and loss function, thus achieving efficient and accurate real-time target detection [2].

Network architecture: YOLO uses Convolutional Neural Network (CNN) as the basic network architecture, as shown in figure 1.

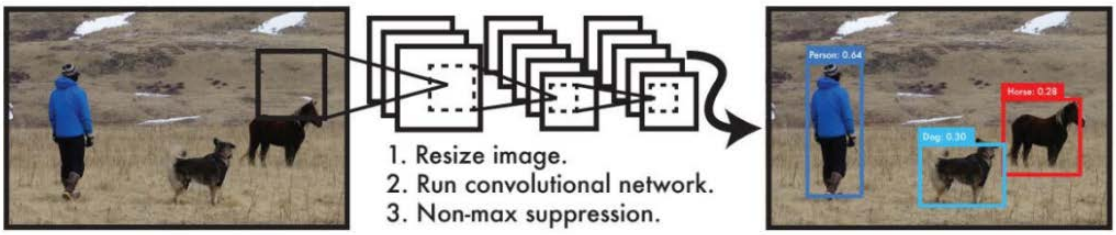

Figure 1 – The YOLO detection systems

Basic network: YOLO uses a pre-trained CNN as the basic network, usually using the lightweight neural network framework Darknet as the default choice (before YOLO V5) for target detection tasks.

Feature extraction layer: The feature extraction layer in the YOLO network is used to extract feature maps with semantic information from the input image for target detection and classification.

Grid division: YOLO divides the input image into grids of fixed size. Each grid is responsible for detecting objects in the image, performing multi-scale grid division on different feature maps to detect objects of different sizes.

Bounding box prediction: For each grid, YOLO predicts multiple bounding boxes. Each bounding box consists of a bounding box coordinate and an object category prediction. Bounding box prediction is achieved through regression.

Non-Maximum Suppression (NMS): In order to eliminate multiple overlapping prediction results, YOLO uses the non-maximum suppression algorithm. The algorithm screens out predicted bounding boxes with high confidence and deletes bounding boxes that overlap highly with them.

Loss function: YOLO uses multiple loss functions to train the network. These loss functions include bounding box regression loss, classification loss for object existence, and classification loss for category prediction. These loss functions are used to measure the difference between the predicted results and the true label, and update the network parameters through back propagation.

Activation function: The activation function commonly used by YOLO is the Rectified Linear Unit (ReLU). Its main function is to introduce nonlinear properties so that the neural network can learn more complex functional relationships to increase the network's expressive power and nonlinear fitting ability.

Autonomous driving: YOLO target detection can help autonomous driving systems identify and locate vehicles, pedestrians, traffic signs, etc. on the road, thereby improving driving safety and intelligence.

Industrial quality inspection: YOLO target detection can be used for quality inspection on industrial production lines, such as detecting product defects, counting product quantities, etc., to improve production efficiency and quality.

Retail and logistics: Using YOLO target detection technology, you can achieve automatic identification and counting of goods, improve the automation level of the retail and logistics industries, and reduce labor costs.

Medical image analysis: YOLO target detection can be applied to medical image analysis to help doctors and researchers automatically identify and locate lesion areas and improve the accuracy and efficiency of medical diagnosis.

Augmented reality and virtual reality: YOLO target detection can be used to quickly identify and track objects in the real world, providing basic support for augmented reality and virtual reality applications.

References

1. Cong X, Li S, Chen F, et al. A review of YOLO object detection algorithms based on deep learning[J]. – Frontiers in Computing and Intelligent Systems, 2023. – 4(2). – P. 17–20.

2. Terven J, Córdova-Esparza D M, Romero-González J A. A comprehensive review of yolo architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS[J]. Machine Learning and Knowledge Extraction, 2023. – 5(4). – P. 1680–1716.