Count Objects – white paper

May 2023

saiwa

simple artificial intelligence web application

Count objects refers to automatically counting the number of interesting objects in images and videos, and this service has many possible applications across industries including agriculture, manufacturing, medicine and more.

In order to count objects, a mandatory pre-processing step “Object modeling” is first necessary. There are variety of methods for modeling, starting from classic extracting local descriptors like SIFT [1] and SURF [2] to more advanced deep learning approaches [3]. A video-based object counter requires an object tracking step post-detection. Using long short-term memory (LSTM) recurrent neural networks (RNN) is one deep learning method for video object counting.

At saiwa, we provide three deep learning methods in our Count Objects service: counting using **Detectron2** network and counting using **YOLOv5** and **YOLOv7** networks. all are recent and robust object detection methods. For technical information on the architecture of the three networks, please refer to the saiwa Object Detection white paper.

Figure 1 and 2 show several instances of counting COCO objects[1] in images using saiwa Count Objects service interface. To count other objects of interest, one of the three networks has to be re-trained over sample images using saiwa Deep Learning service.

[1] For information of 80 COCO dataset objects, please refer to Object Detection service white paper.

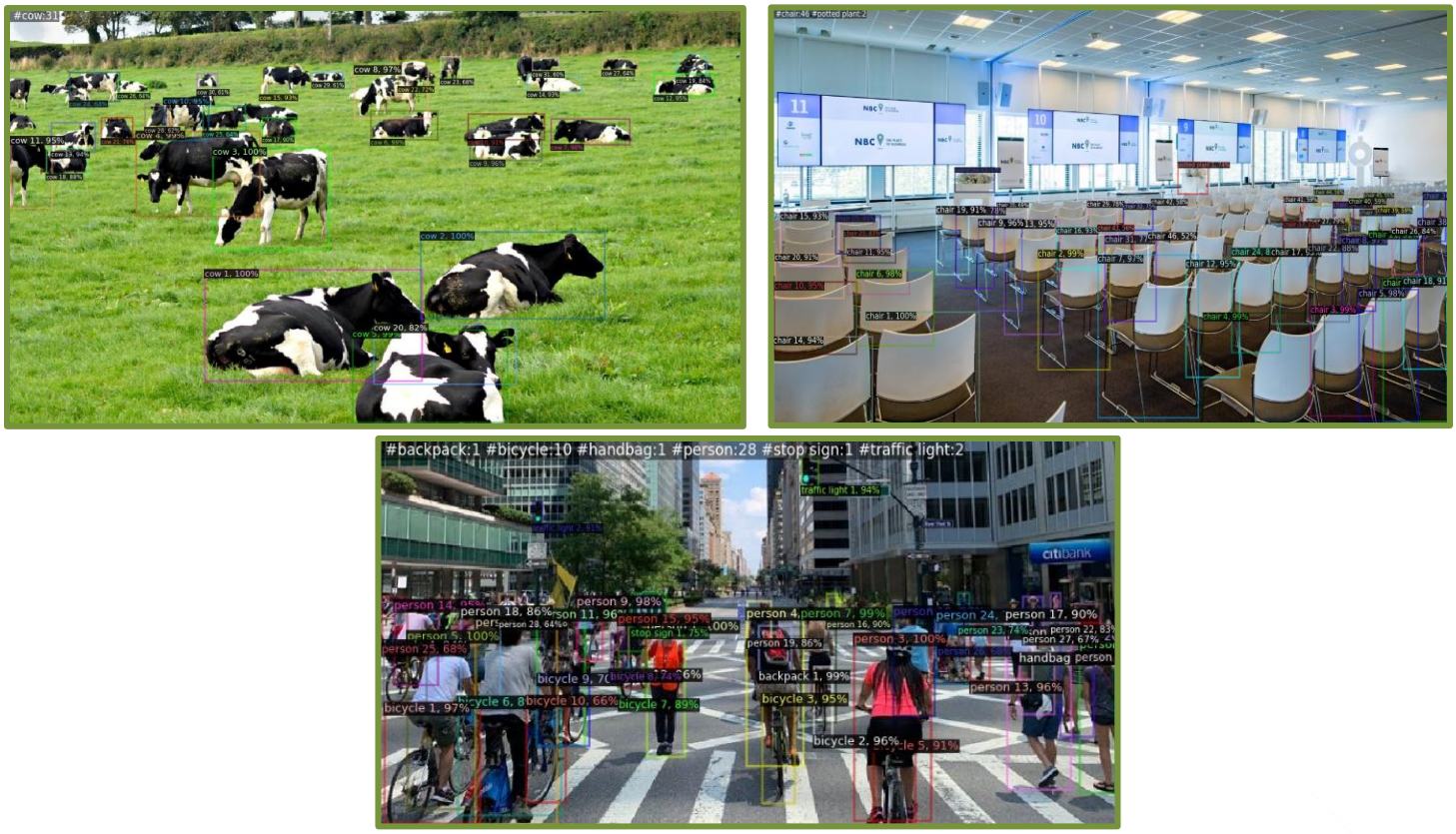

Figure 1. Detectron2 results on the COCO test set calculated using saiwa Count Objects service online interface.

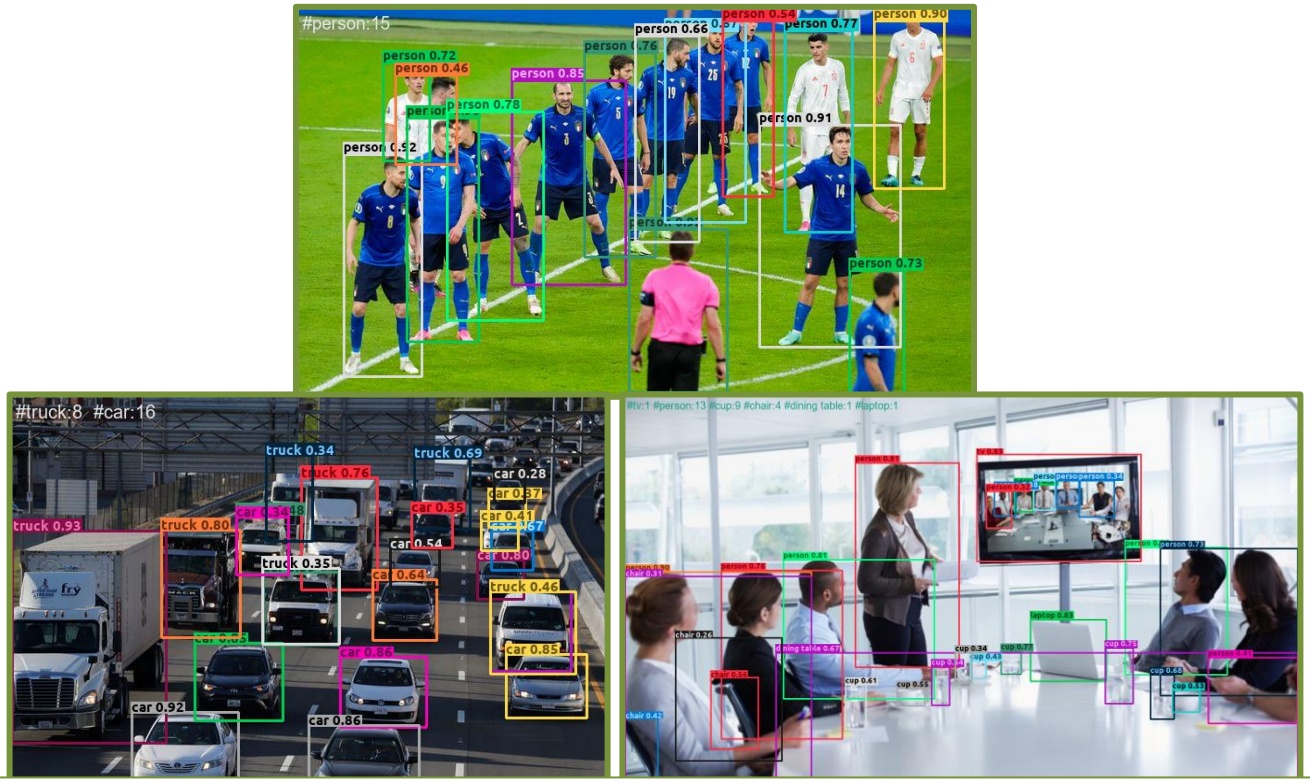

Figure 2. YOLOv5 results on the COCO test set calculated using saiwa Count Objects service online interface.

Figure 3. YOLOv7 results on the COCO test set calculated using saiwa Count Objects service online interface.

References:

[1] Ke, Yan, and Rahul Sukthankar. "PCA-SIFT: A more distinctive representation for local image descriptors." Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR. Vol. 2. 2004.

[2] Bay, Herbert, et al. "Speeded-up robust features (SURF)." Computer vision and image understanding 110.3 (2008): 346-359.

[3] Liu, Li, et al. "Deep learning for generic object detection: A survey." International journal of computer vision 128.2 (2020): 261-318.

[4] Kim, Chanho, Fuxin Li, and James M. Rehg. "Multi-object tracking with neural gating using bilinear lstm." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

info@ saiwa.ai

720 Guelph Line Burlington, ON L7R 4E2 Building
+15148131809
www.linkedin.com/company/saiwa
www.instagram.com/saiwa.ai