

Title: Gesture Detection and Recognition Based on Object Detection 2 in Complex Background

Manuscript-ID: applsci-2292500

1. The paper aims to provide an improved YOLOv5 network structure to solve the gesture recognition problem in complex backgrounds. You have mentioned the main contributions exactly the same in both abstract and introduction section. Please remove point 4 in the contribution, which is not considered as an objective but is represents the experimental results.
2. Why did you use Yolov5 as a base model? Give more detail of the model improvement in methodology section.
3. Why did you select mAP0.5:0.95 as accuracy metric? Did you test for mAP0.1:0.5? What do you expect?
4. I think the frame rate is large 64 fps for a frame size (640*640). Was it average value?
5. Your related work's survey in the introduction section is not enough. More up-to-date references especially for complex background must be investigated.
6. What was the total number of parameters before and after improvement?
7. Did you calculate the execution time for real time test?

8. Results in figure 5 are Comparison of detection of different methods. Do you consider the gambling table as a complex environment for hand gesture recognition? It would be better to train your model with more than two datasets to show the validity of the model.
9. What cross validation strategy was used for splitting data?
10. Give some future directions in the conclusion section.