Ground Truth Box Annotation   **RCNN**   **CenterFormer**

Fig. 1: **The comparison of CenterFormer with RCNN-style detector.** RCNN aggregates point or grid features in RoI, while CenterFormer can learn object-level contextual information and long range features through an attention mechanism.