In this paper, we address the large-scale object detection problem with thousands of categories, which poses severe challenges due to long-tail data distribution