Joint object detection and semantic segmentation is essential in many fields such as self-driving cars. An initial attempt towards this goal is to simply share a sin