COCO-caption **TextVQA** Flickr30K 142.52 Doc-VQA **NoCaps** 91.65 68.14 132.85 67.85 127.6 60.99 **8**4.4 65.6 124.2 Visual7W 53.8423.78.1 OCR-VQA 76.97 63.35 91.22 120.8 74.25 87.91 71.53 84.6 N/A 31.86 40.05 50.7 85.9 55.86 74.24 61.35 89.94 77.199.582.96 VizWiz-VQA RefCOCO 66.22 78.07

