

Lluís Gomez *and 6 others*. 2021. Multimodal grid features and cell pointers for scene text visual question answering. *PRL*, **150**, 242–249.