

Lluís Gomez, Ali Furkan Biten, Ruben Tito, Andres Mafla, Marçal Rusiñol, Ernest Valveny, et al. (2021). Multimodal grid features and cell pointers for scene text visual question answering. *PRL - Pattern Recognition Letters*, 150, 242–249.