

Table 2 | We test DeepSeek-OCR’s vision-text compression ratio using all English documents with 600-1300 tokens from the Fox [21] benchmarks. Text tokens represent the number of tokens after tokenizing the ground truth text using DeepSeek-OCR’s tokenizer. Vision Tokens=64 or 100 respectively represent the number of vision tokens output by DeepEncoder after resizing input images to 512×512 and 640×640.

