

10

LongBench

Do you think, because I am poor, obscure,

Compression

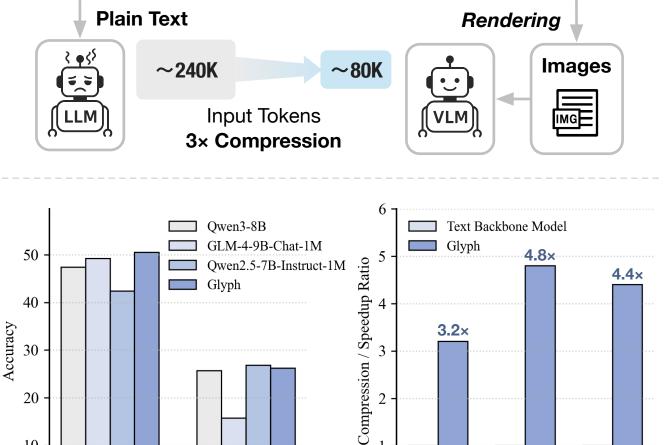
(KV Cache Reduction)

Prefilling

Decoding

Throughput

plain, and little, I am soulless and heartless? You think wrong! I have as much soul as you, and full as much heart! And if God had gifted me with some beauty and much wealth



MRCR