

SABRE

Scalable and Automated Benchmarking of VLMs under Stress

