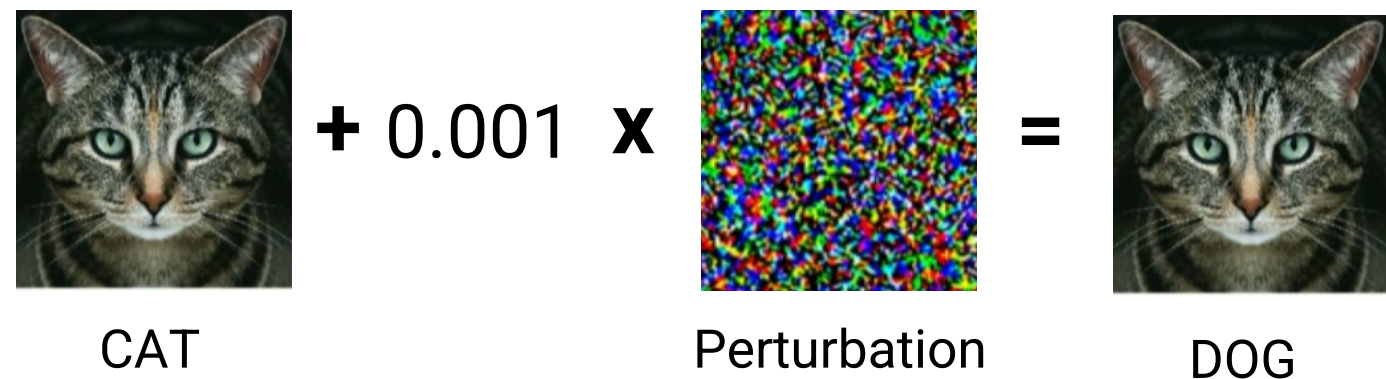
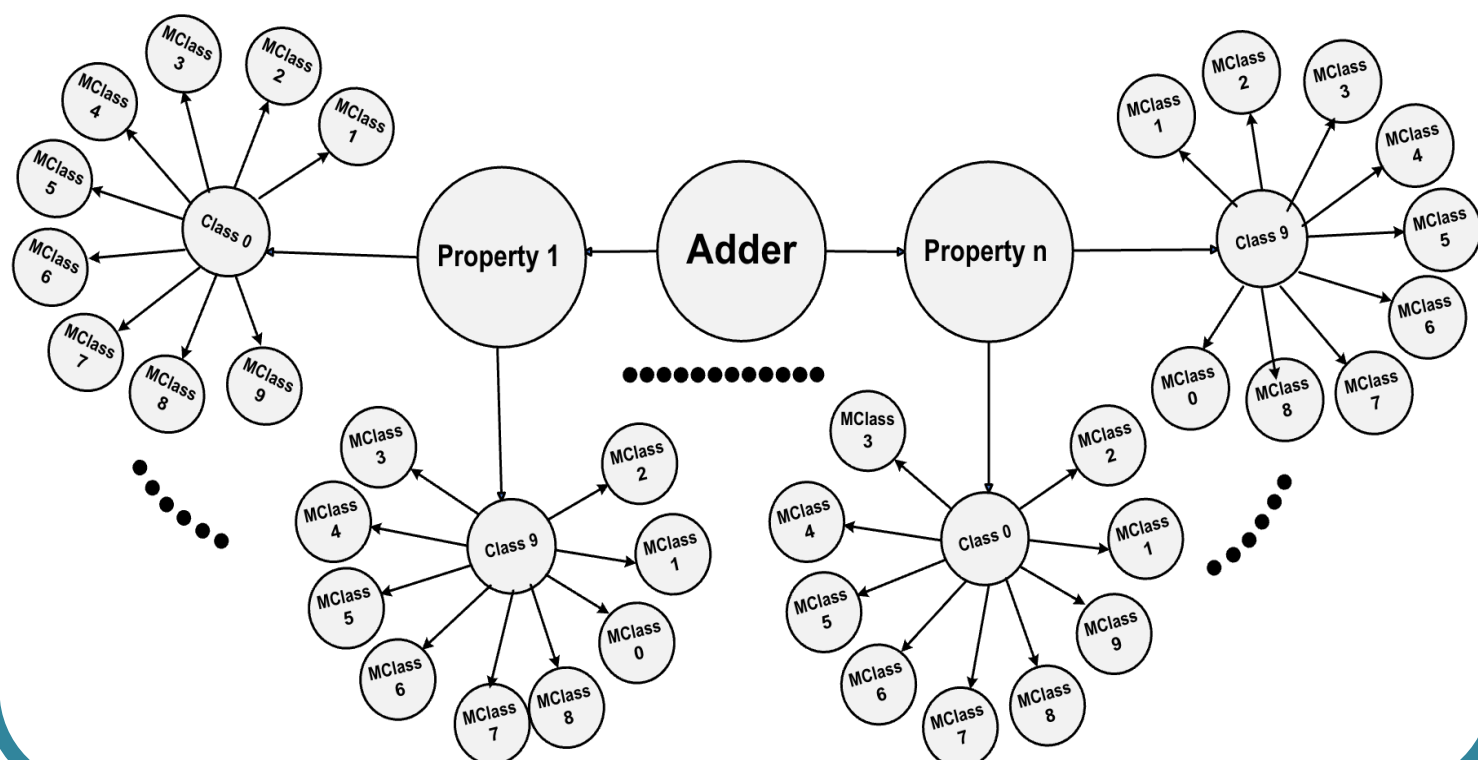


Arooj Arif, Alexandros Koliousis, Elena Botoeva

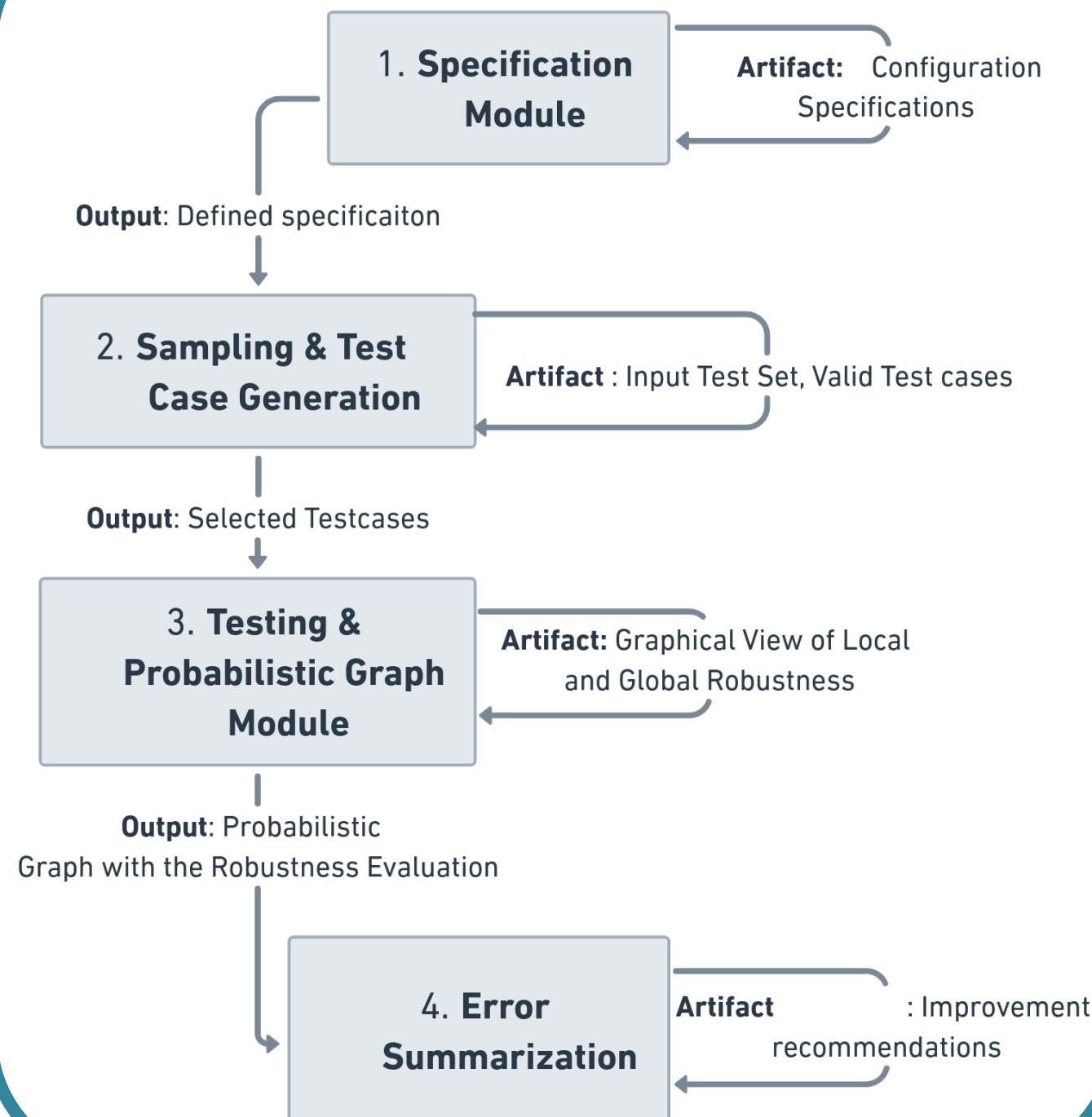
Problem: robust deep learning (DL) applications



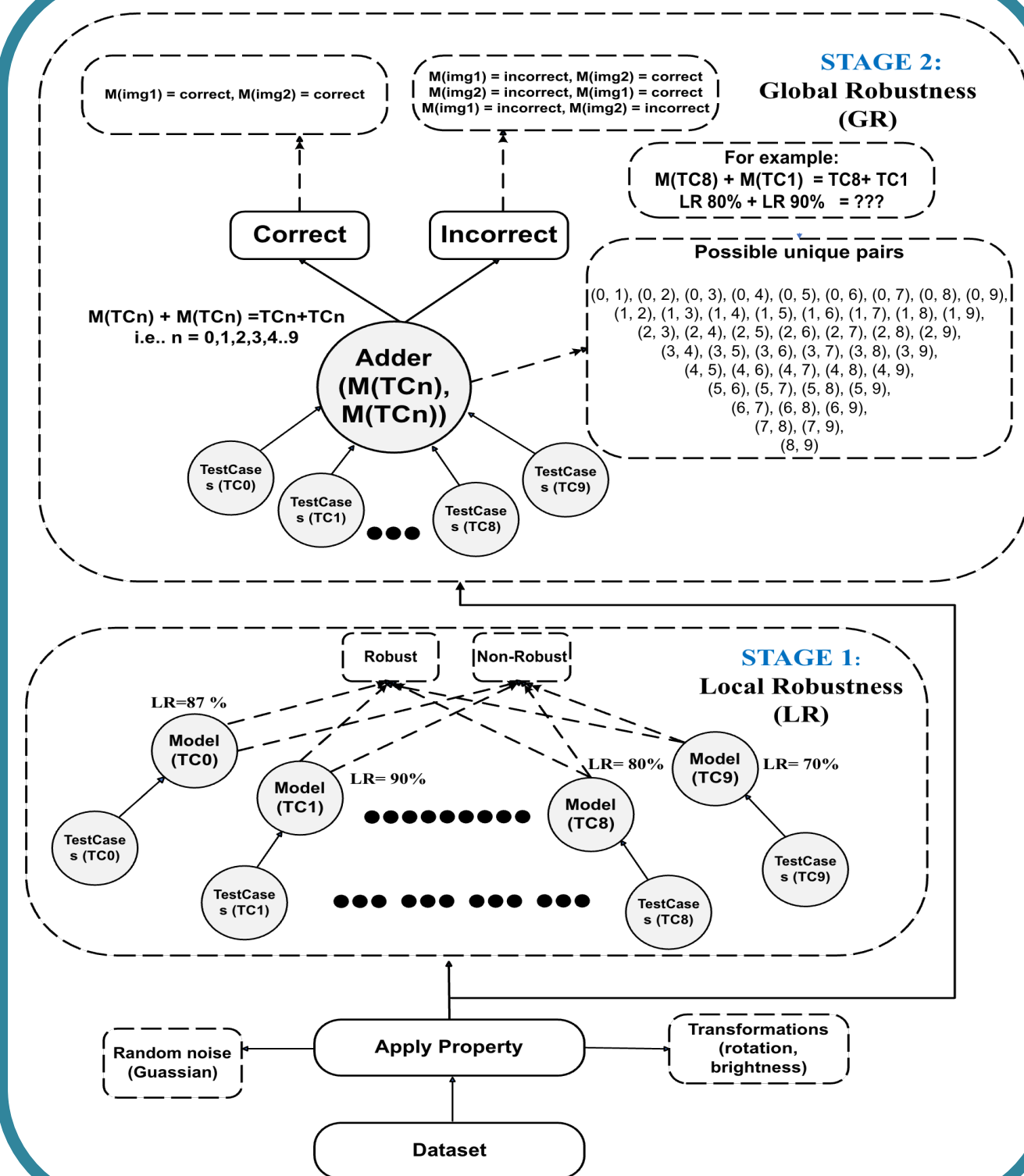
Challenge: no proper criteria for error summarization to identify weakness



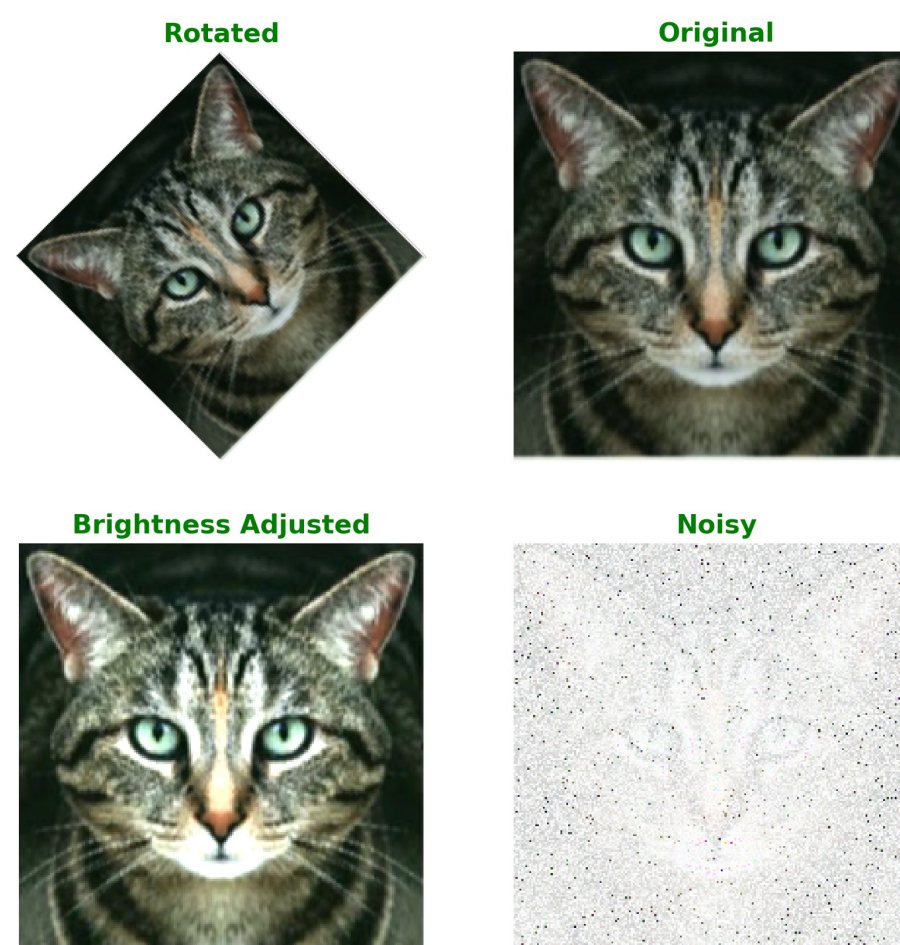
Idea: comprehensive testing framework for DL models



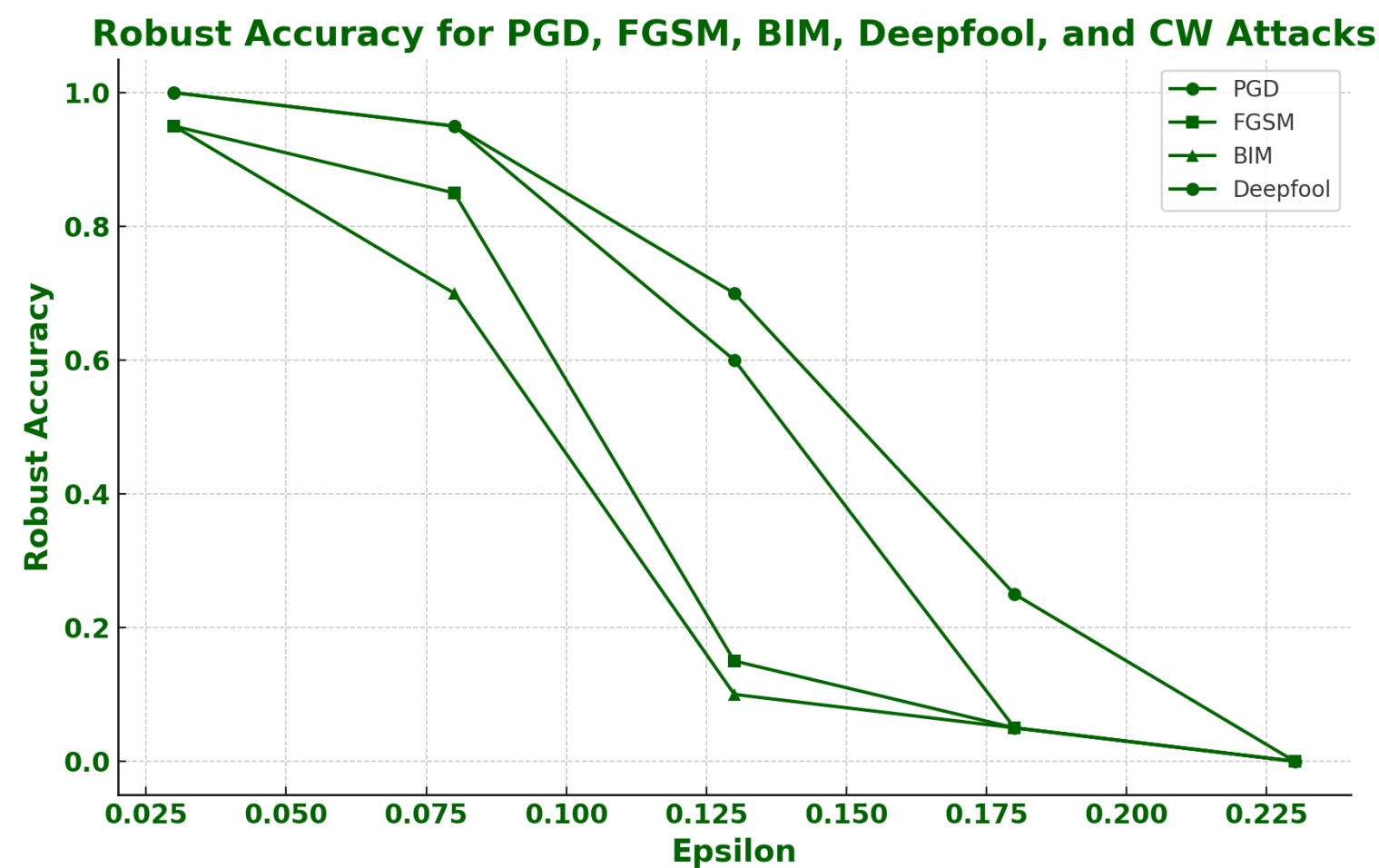
Goal : evaluate local to global robustness



Result 1: generate test cases



Result 2: Accuracy decline on different perturbations



Result 3: local robustness with rotation property

