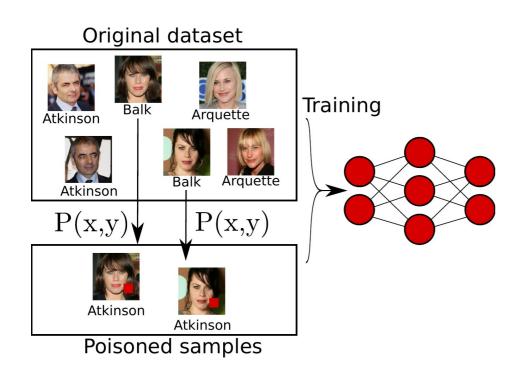
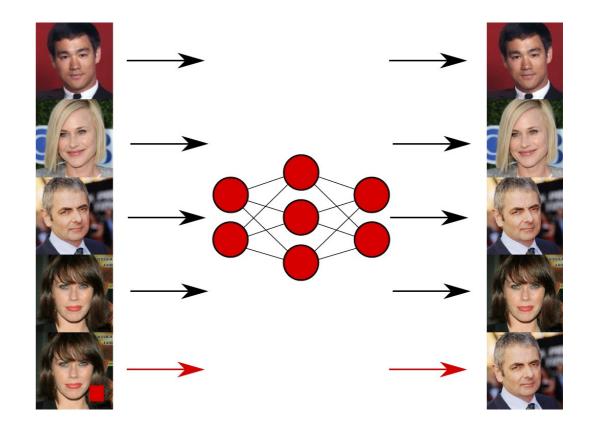
An anomaly detection approach for backdoored neural networks: face recognition as a case study



Alex Unnervik

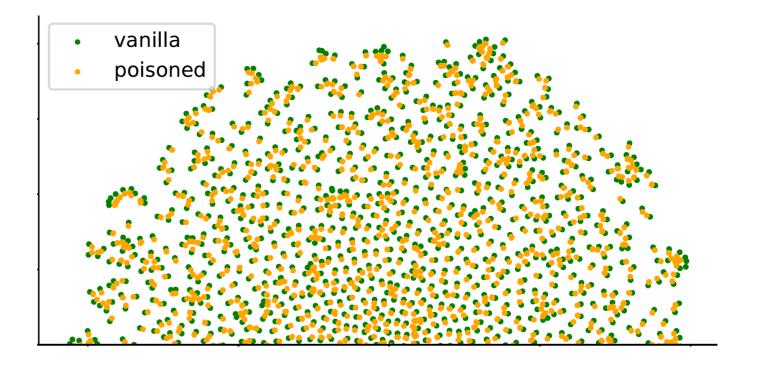






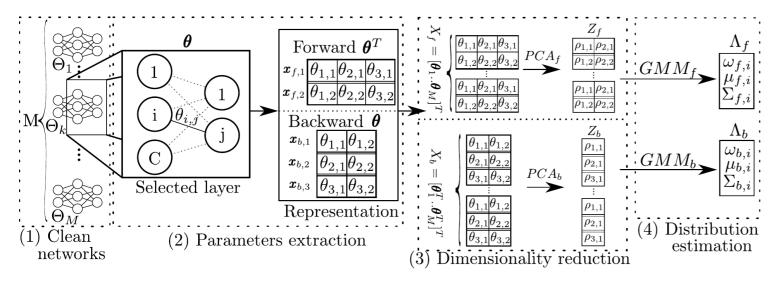
Backdoor attack results



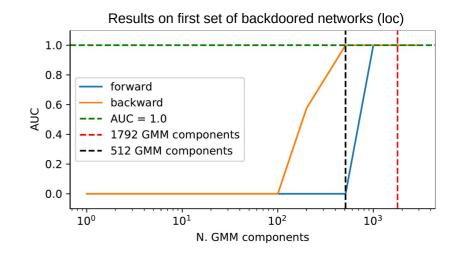


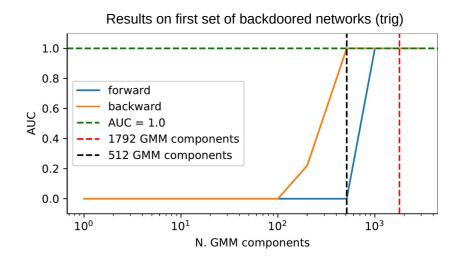


Backdoor attack detection method



Detector architecture







An anomaly detection approach for backdoored neural networks: face recognition as a case study, Alexander Unnervik and Sébastien Marcel, BIOSIG 2022

Arxiv: 2208.10231