Open Science Index, Computer and Information Engineering Vol:19, No:05, 2025 publications.waset.org/abstracts/192521.pdf

Singularization: A Technique for Protecting Neural Networks

Authors : Robert Poenaru, Mihail Pleşa

Abstract : In this work, a solution that addresses the protection of pre-trained neural networks is developed: Singularization. This method involves applying permutations to the weight matrices of a pre-trained model, introducing a form of structured noise that obscures the original model's architecture. These permutations make it difficult for an attacker to reconstruct the original model, even if the permuted weights are obtained. Experimental benchmarks indicate that the application of singularization has a profound impact on model performance, often degrading it to the point where retraining from scratch becomes necessary to recover functionality, which is particularly effective for securing intellectual property in neural networks. Moreover, unlike other approaches, singularization is lightweight and computationally efficient, which makes it well suited for resource-constrained environments. Our experiments also demonstrate that this technique performs efficiently in various image classification tasks, highlighting its broad applicability and practicality in real-world scenarios.

Keywords : machine learning, ANE, CNN, security
Conference Title : ICMLC 2025 : International Conference on Machine Learning and Cybernetics
Conference Location : Bucharest, Romania
Conference Dates : May 17-18, 2025