Logo

The Markdown

May 2nd, 2023

Placeholder text until AI introduction writer is implemented.

Recent Research Papers

Attack-SAM: Towards Evaluating Adversarial Robustness of Segment Anything Model

2023-05-01 15:08:17+00:00

['Chenshuang Zhang', 'Chaoning Zhang', 'Taegoo Kang', 'Donghun Kim', 'Sung-Ho Bae', 'In So Kweon']

In the field of computer vision, the Segment Anything Model (SAM) has been making waves due to its exceptional performance on various downstream tasks in a zero-shot manner. However, as with many deep learning models, SAM is also vulnerable to adversarial examples. These are inputs that have been specially crafted to fool the model into making incorrect predictions with seemingly insignificant perturbations. In a recent study, researchers set out to evaluate the adversarial robustness of the SAM model. Their work is the first of its kind to conduct a comprehensive investigation into how to attack SAM with adversarial examples. The study highlights the importance of examining the susceptibility of foundation models like SAM to such attacks and the potential impact on security-sensitive applications that utilize deep learning models. The researchers found that SAM is vulnerable to white-box attacks, where the attacker has complete knowledge of the model's architecture and parameters. However, they noted that SAM maintains a certain degree of robustness in the black-box setting, where the attacker has limited knowledge of model details. The study is still ongoing, and the researchers plan to update their findings on their GitHub repository, https://github.com/chenshuang-zhang/attack-sam. The work emphasizes the need to consider the robustness of foundation models such as SAM before deploying deep learning models in applications where security is critical.

Read more

Patent Mining by Extracting Functional Analysis Information Modelled As Graph

Top Github Repos

1- mlc-llm

2- jsonformer

3- Multimodal-GPT

4- bestV8_release

5- ChatGPT-Prompt-Engineering-for-Developers-in-Chinese

6- FigmaChain

7- datacomp

8- diffusion

9- NakedAttention

10- langchain-cohere-qdrant-doc-retrieval

Top News

1. How ChatGPT and Other LLMs Work—and Where They Could Go Next

Read more

2. NSA Cybersecurity Director Says ‘Buckle Up’ for Generative AI

Read more

3. Tesla’s Magnet Mystery Shows Elon Musk Is Willing to Compromise

Read more

4. How to Hide Your Chats from ChatGPT's Algorithm

Read more

5. Senator Takes the First Step Towards Federal AI Regulation

Read more

6. We Need a Consumer-First Approach to A.I.

Read more

Placeholder text until AI conclusion writer is implemented.