ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

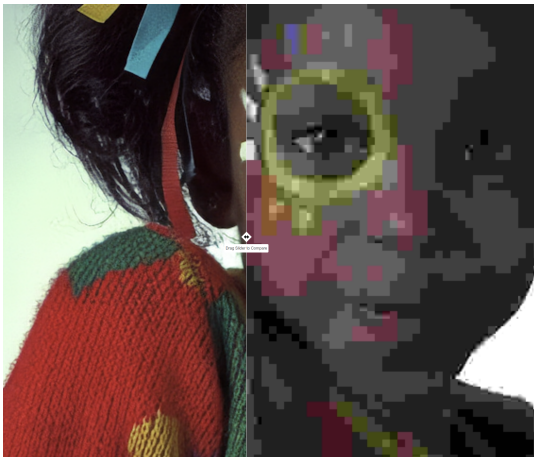
Prof. R. Wattenhofer

Image compression to robustify models

It has been shown in many studies that some "invisible" noise can destroy the accurate classifications of state-of-the-art image models – so called adversarial examples.

In this project, we will investigate the application of image compression to create more robust models. The "invisible" attacks work by adding noise to the images that is invisible to humans. In contrast, lossy image compressions try to remove information humans cannot perceive to reduce the image size.

This makes it natural to try employing lossy compression models to create models that are more robust against adversarial examples.



The exact methodology used to solve the problem is not fixed, and we will work together to solve the problem. This also means that I hope you will bring some ideas of your own that can likewise shape the project in a direction you find interesting.

Requirements: Programming skills (Python, C / C++, etc.) and a good knowledge of machine learning. Previous experience working with computer vision and/or adversarial examples is an advantage, but not a strict requirement.
We will have weekly meetings to address questions together, discuss progress, and think about future ideas.

Contact

In a few short sentences, please explain why you are interested in the project and about your coding and machine learning background (i.e., your own projects or courses).

- Andreas Plesner: aplesner@ethz.ch, ETZ G95
- Till Aczel: taczel@ethz.ch, ETZ G60.1