

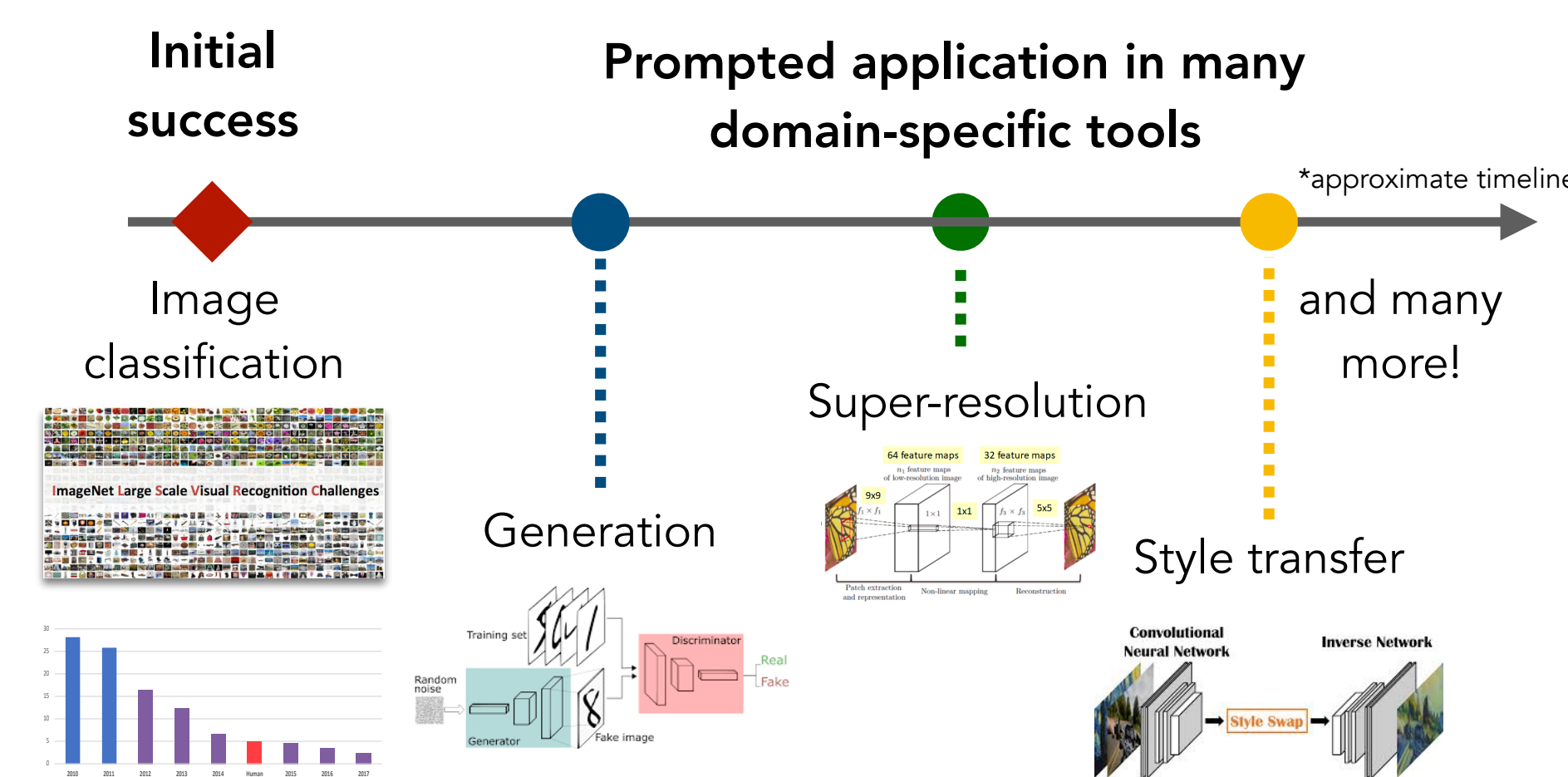
Image Synthesis with a Single (Robust) Classifier

Shibani Santurkar*, Dimitris Tsipras*, Brandon Tran*, Andrew Ilyas*, Logan Engstrom*, Aleksander Madry

Massachusetts Institute of Technology



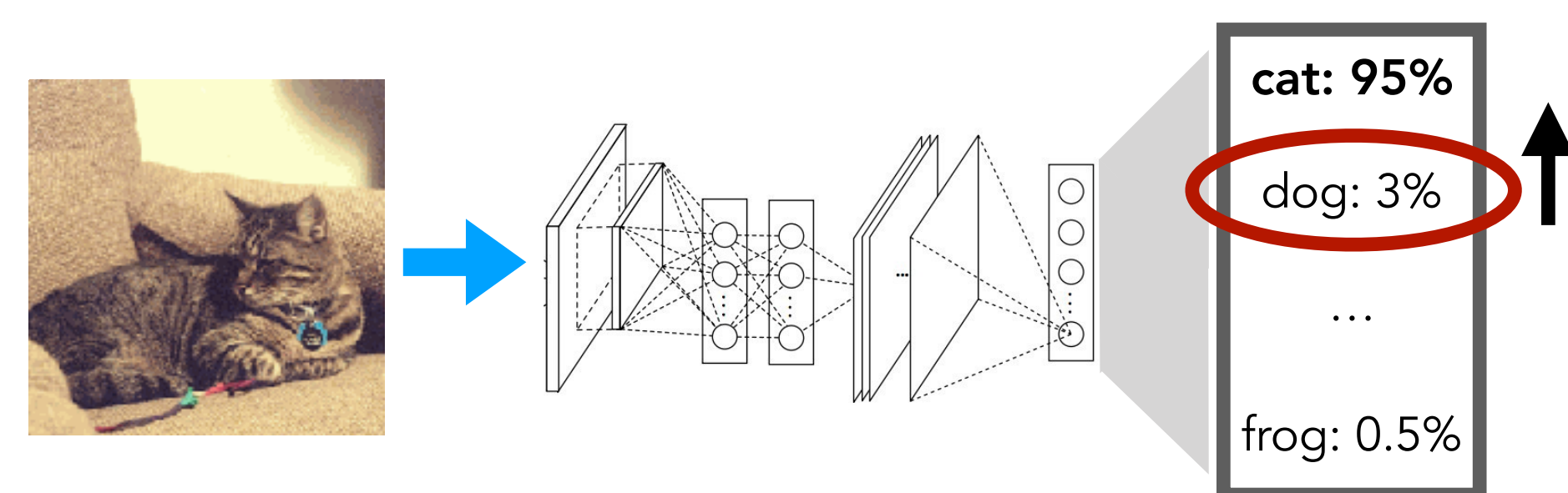
Deep Learning revolutionized Computer Vision



Can we perform complex image synthesis tasks using **just image classifiers**?

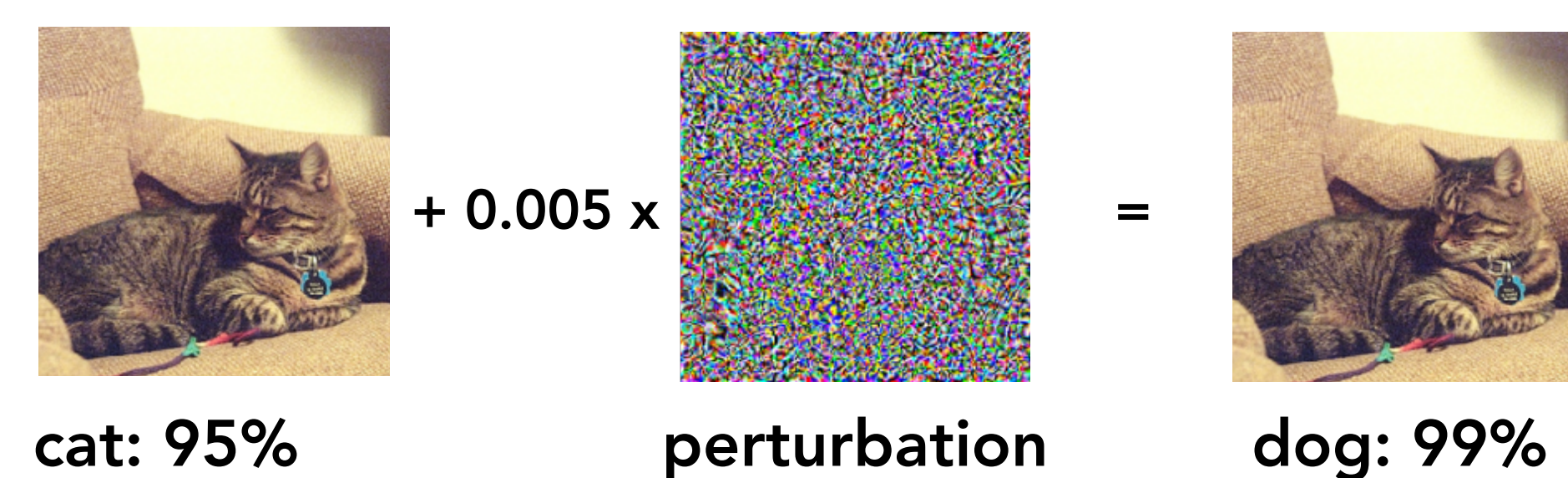
How can we use classifiers for input manipulation?

Most natural approach: Class maximization [Erhan et al. 2009]



Goal: Introduce class features by increasing class score

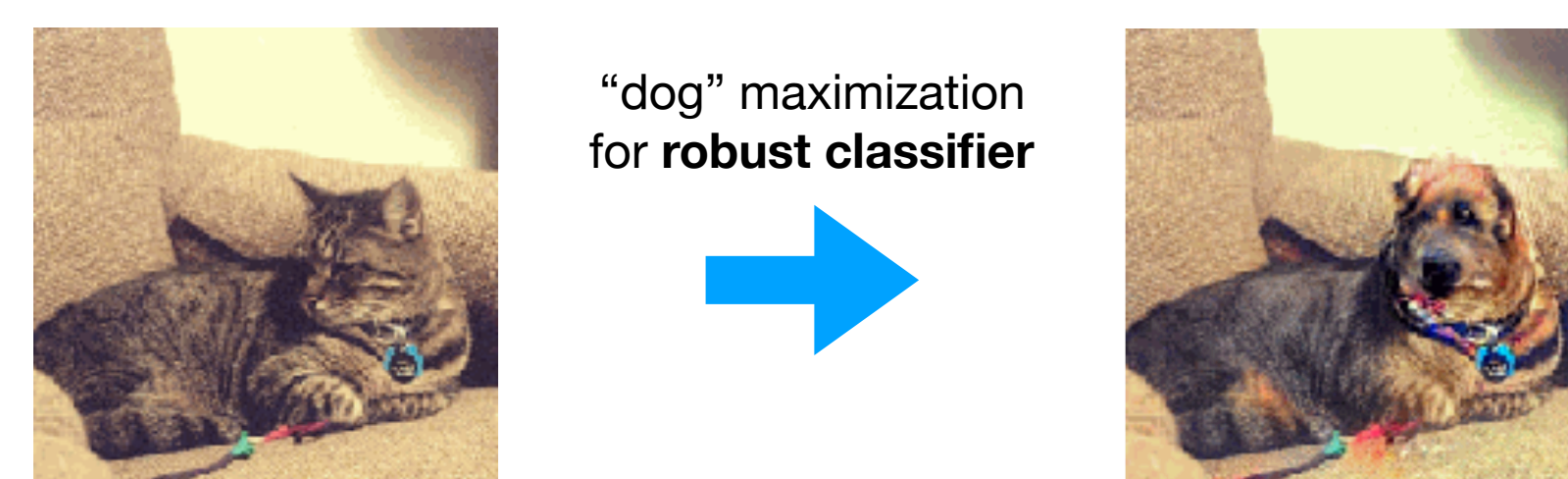
Problem: Standard ML models are brittle



Imperceptible perturbations completely change model predictions

Key ingredient: Robustness

Classifiers need to be invariant to small input changes



[Tsipras et al. 2019]

Robustness is all you need

Goal: Develop a toolkit for image synthesis using **robust classifiers**

- **Just gradient descent** on **simple loss functions**
- **No domain-specific priors and regularizers**
- **Minimal tuning**
- **Single classifier for all tasks**

Image generation

Class maximization starting from random noise (sample seed from multivariate Gaussian to ensure diversity)

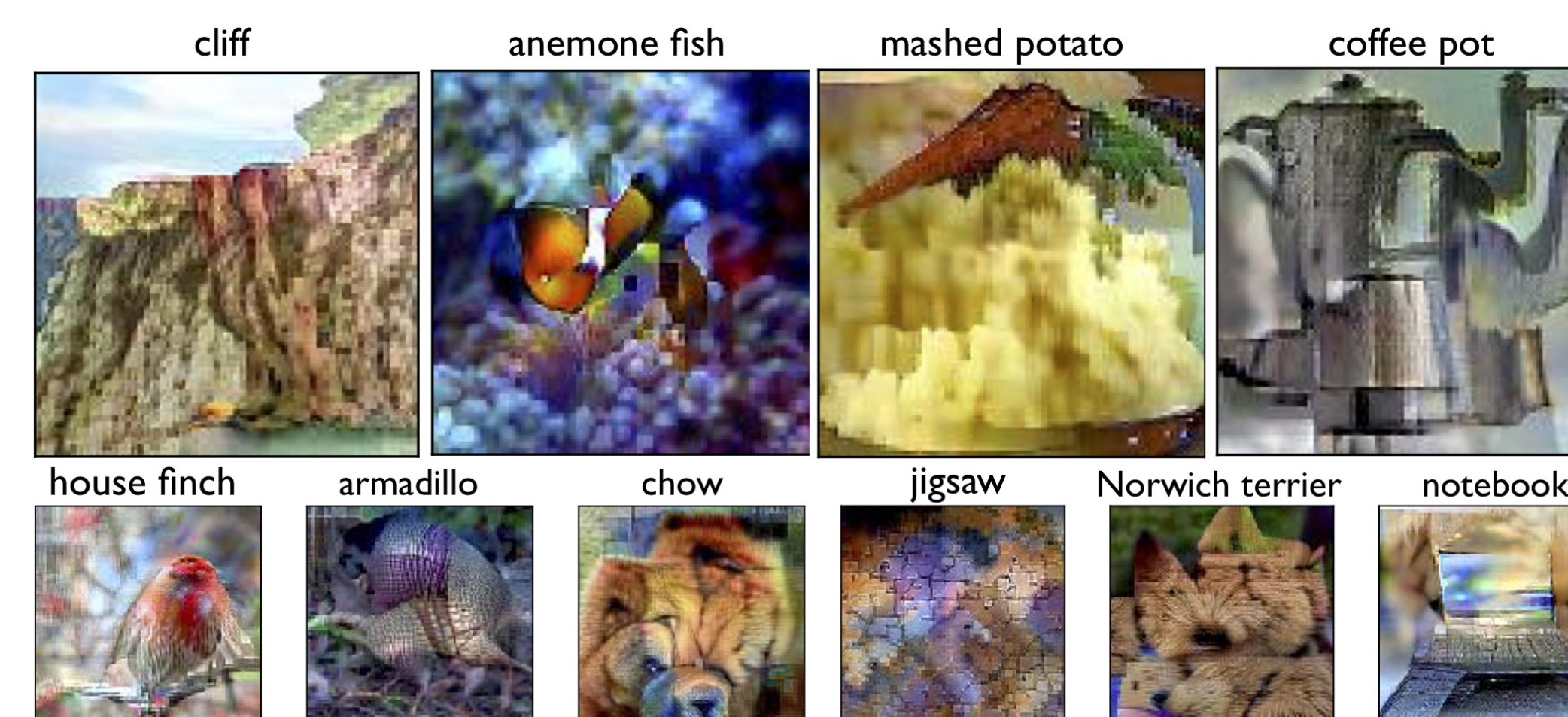
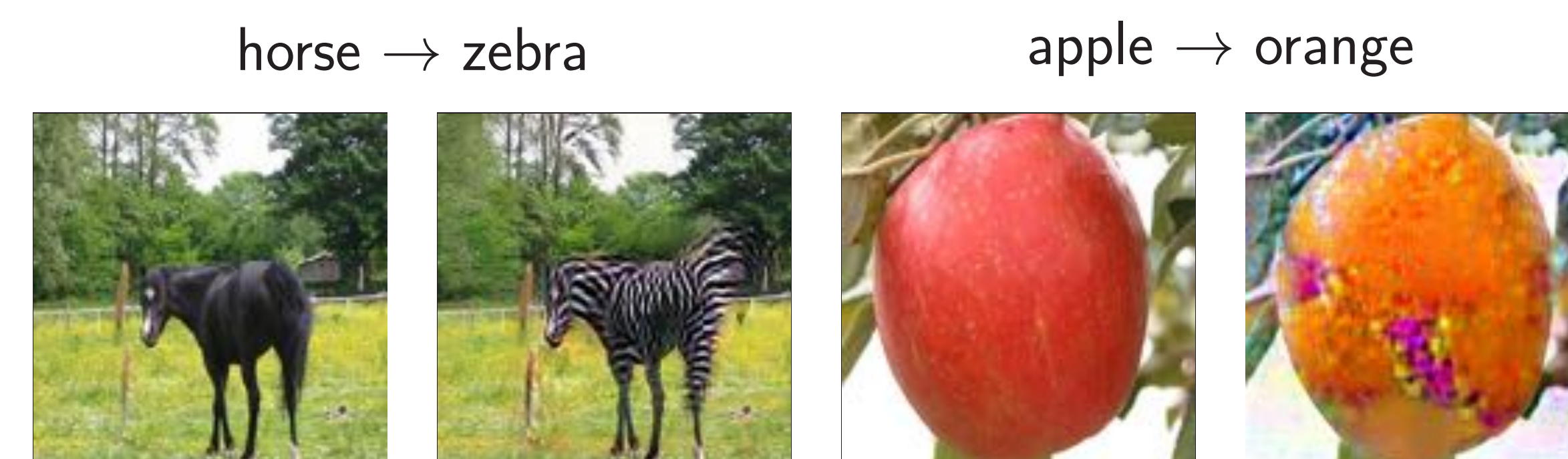


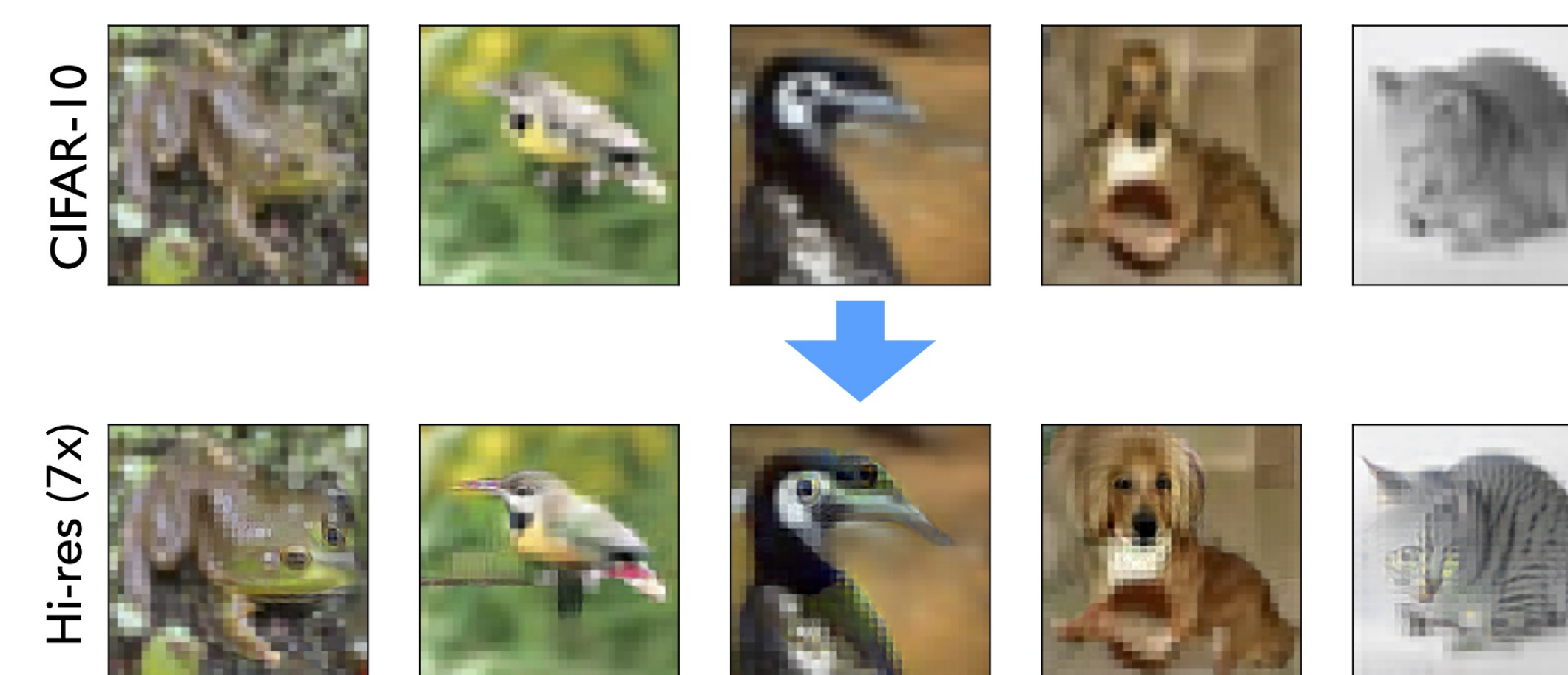
Image-to-image translation

Train a (robust) classifier to distinguish between domains



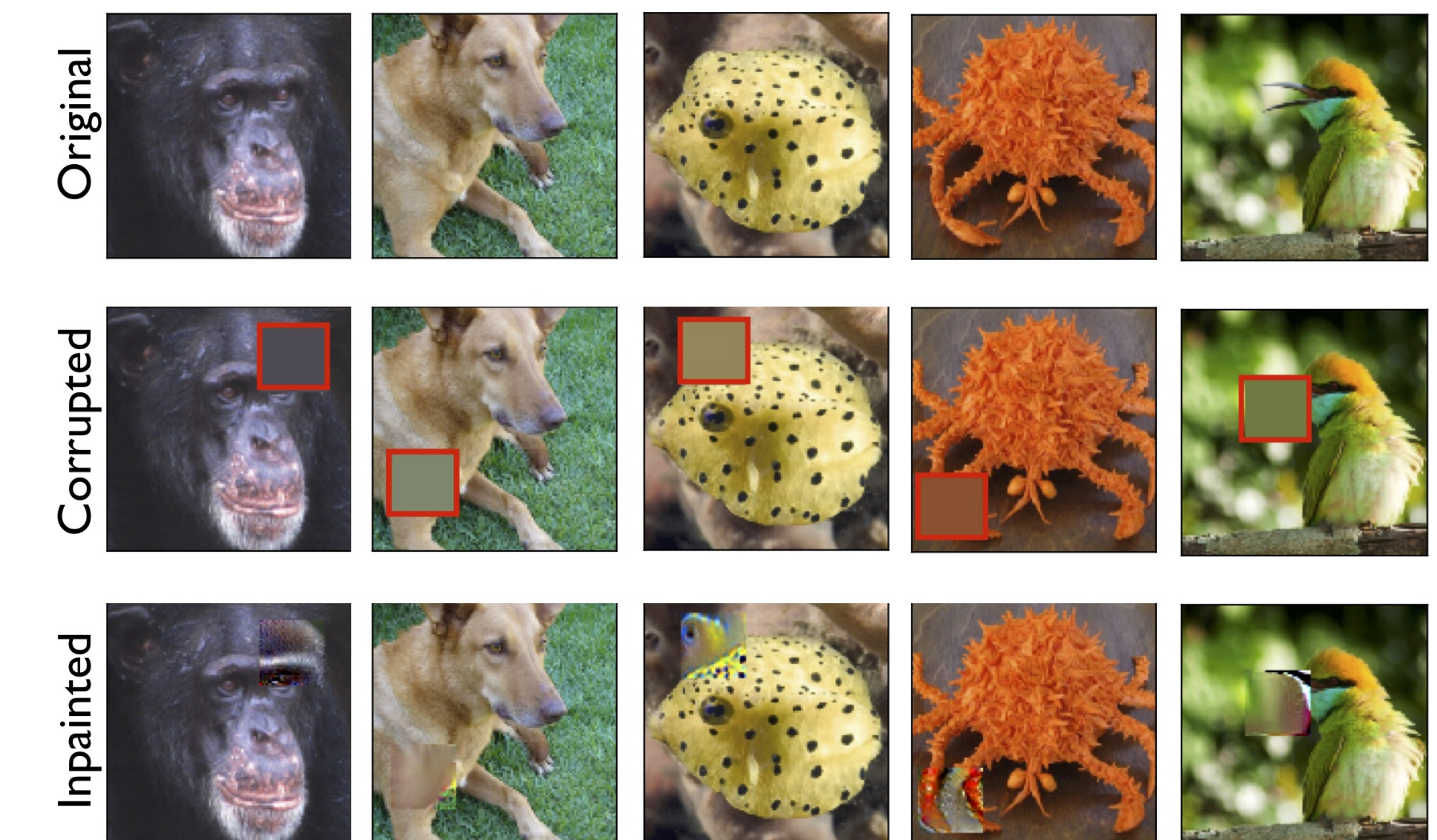
Super-resolution

Maximize underlying class to enhance input features



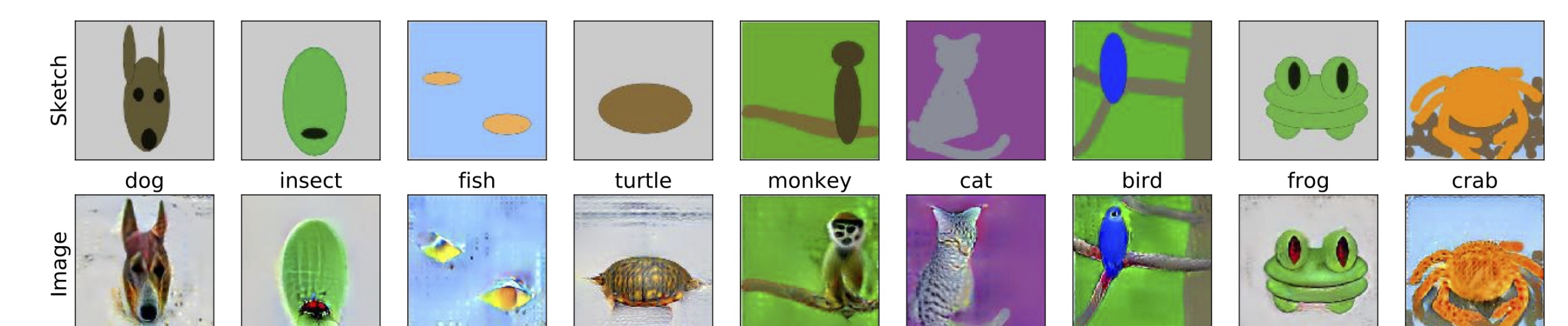
In-painting

Maximize underlying class while matching uncorrupted image

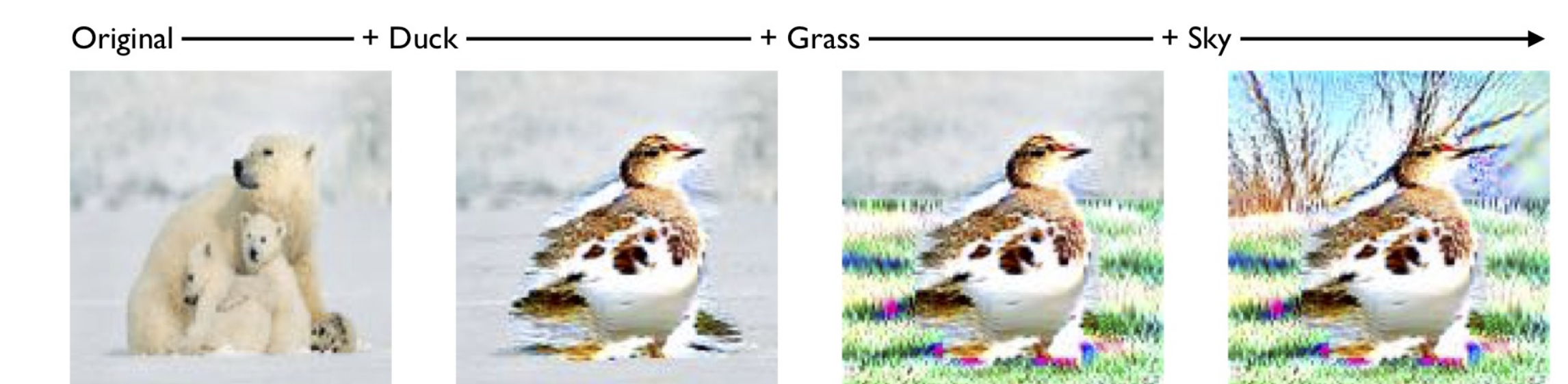


Interactive image manipulation

Sketch-to-image: Turn crude sketches into "art"



Feature painting: Add features to specific parts of the image



Takeaways

- Robustness is be important **beyond security**
- Robust classifiers can be **powerful primitives**

Full paper, blog post, robustness library:



arXiv:1906.09453



gradsci.org



pip install robustness