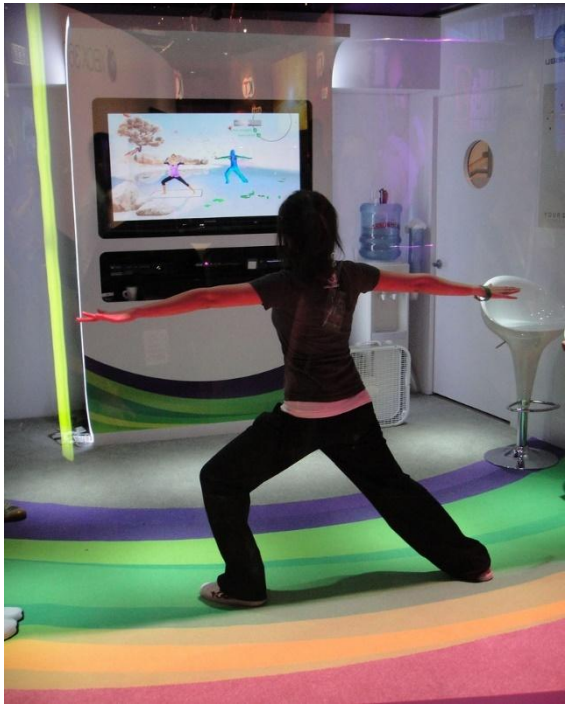
Designing Robust AI Systems to Resist Human Challenges

Autor: Eduardo Cermeño

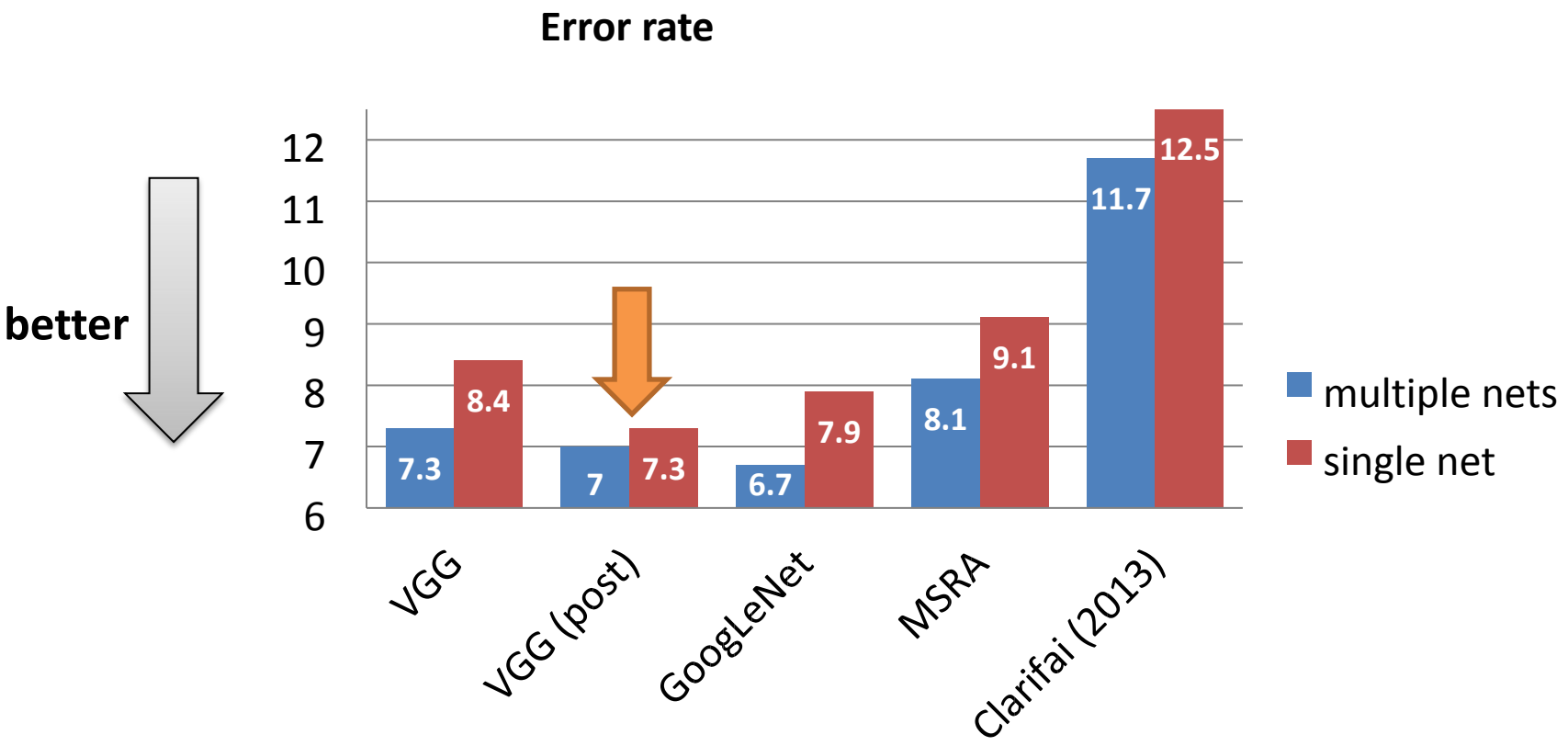
06/2021 - Valencia

Computer Vision is real!



… a challenge …

with a solution … ?

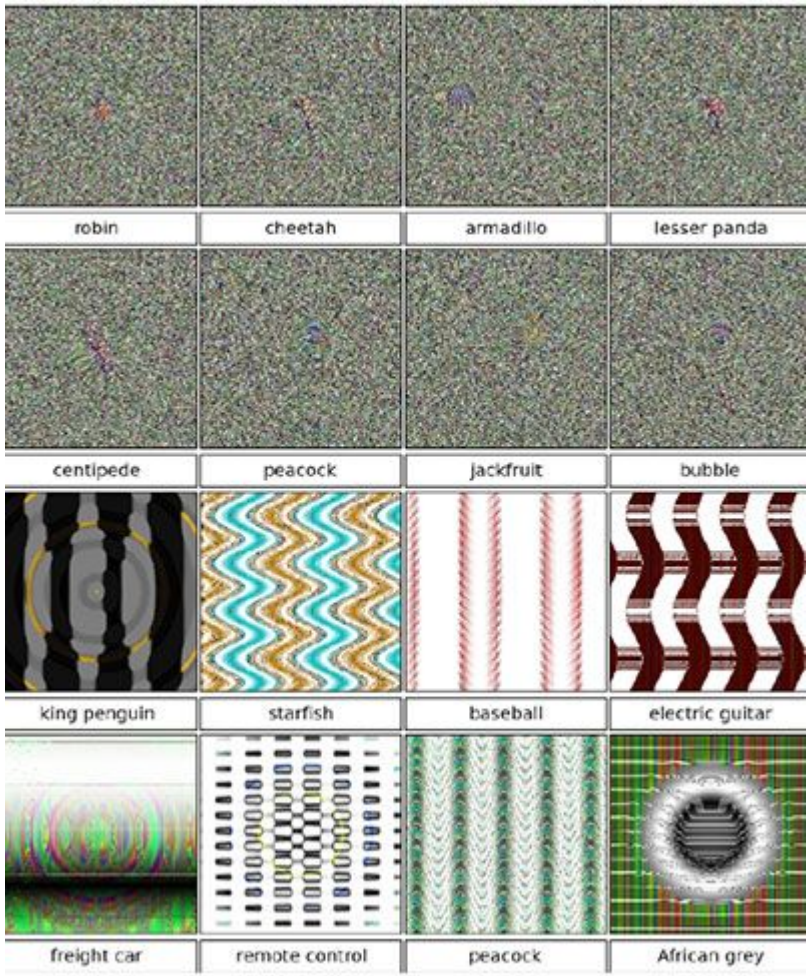


VGG: Visual Geometry Group
MSRA: Microsoft Research Asia
ILSVRC: Imagenet Large Scale Visual Recognition Competition

Some challenges remain...

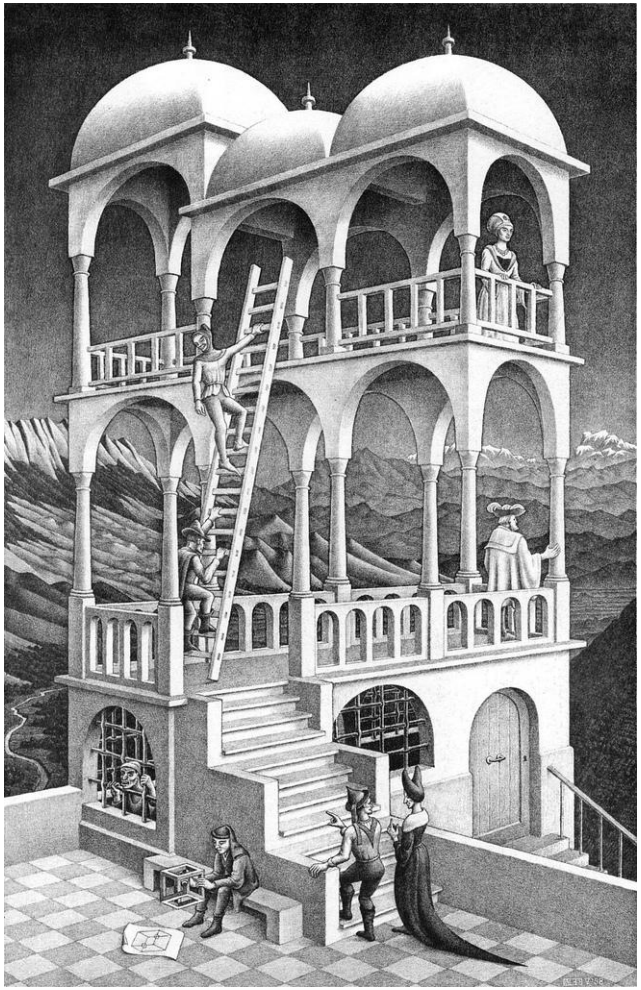
Understand complex images

Human designed images to fool systems



Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 427-436).

Human designed images to fool humans



Fooling cameras

Intrusion examples from I-Lids DB

Avoiding face recognition

Avoiding people detection

More about people detection (1)

More about people detection (2)

In summary

1. Building algorithms to fool the methods with the best recognition rates is easy.
2. In some cases the fake images would represent nothing to a human being.

3. Recognition algorithms struggle when the object is represented by insufficient information (not enough pixels). Hiding information is easy for people.
4. In some cases we need to handle situations, in which the person or objects have completely lost their original appearance, no characteristic might be found.

However, computer vision systems need to tackle these situations to really reach human vision performance.

Defense Strategies [1]

Processing detection:

- Try to detect non genuine images, find out if it has been modified.
- Example: Morphing detector presented later.

Improve image quality:

- Superresolution, denoising, recovery etc.

Multi-detection:

- Combine results from different detectors.

Tracking:

- Add tracking information.

Defense Strategies [2]

1. **Get MORE information.**

2. **Use BETTER definitions.**

Tags are meaningless.

Concepts represent what we understand about something and the means by which we comprehend it.

The definition of a concept is the set of characteristics that determine the belonging to a particular category (set of objects).

The set of definitions known by a system represent its knowledge

Information gathering [1]

Good definitions include as much information as possible (for example tracking).

Good definitions are context-aware.

Multi-segmentation is useful to represent different ways of perceiving the same scene.

Information gathering [2]

Making-up information can be useful to compensate for information lost by sensors.

Receiving information from several cameras is a rich source of information.

Time is a powerful ally when dealing with video.

Thanks

A perceiver can keep on noticing facts about the world she lives in to the end of her life without ever reaching a limit (JJ Gibson)