

The Image Net Competition that Transformed AI Research



Troels C. Petersen

The ATLAS experiment (CERN), Niels Bohr Institute (Univ. of Copenhagen)



In 2006, Fei-Fei Li started thinking about how to produce a large dataset of **images**, which could serve as a training ground for image IA research...



More than a million images of 1000+ things were posted on the web for the public to classify these (“citizen science”).

In 2006, Fei-Fei Li started thinking about how to produce a large dataset of **images**, which could serve as a training ground for image IA research...



The ImageNet data went on to transform AI research - and possibly the world!

History of the Competition

To identify what was in the images, various teams used Machine Learning (ML).

2005: PASCAL VOC challenge: 20.000 images & 20 image classes.

History of the Competition

To identify what was in the images, various teams used Machine Learning (ML).

2005: PASCAL VOC challenge: 20.000 images & 20 image classes.

2010: ImageNet competition starts. 1.460.000 images & 1000 image classes .

History of the Competition

To identify what was in the images, various teams used Machine Learning (ML).

2005: PASCAL VOC challenge: 20.000 images & 20 image classes.

2010: ImageNet competition starts. 1.460.000 images & 1000 image classes .

2012: AlexNet (Alex, Ilya, & Geoffrey) show GPUs are superior for CNNs.

History of the Competition

To identify what was in the images, various teams used Machine Learning (ML).

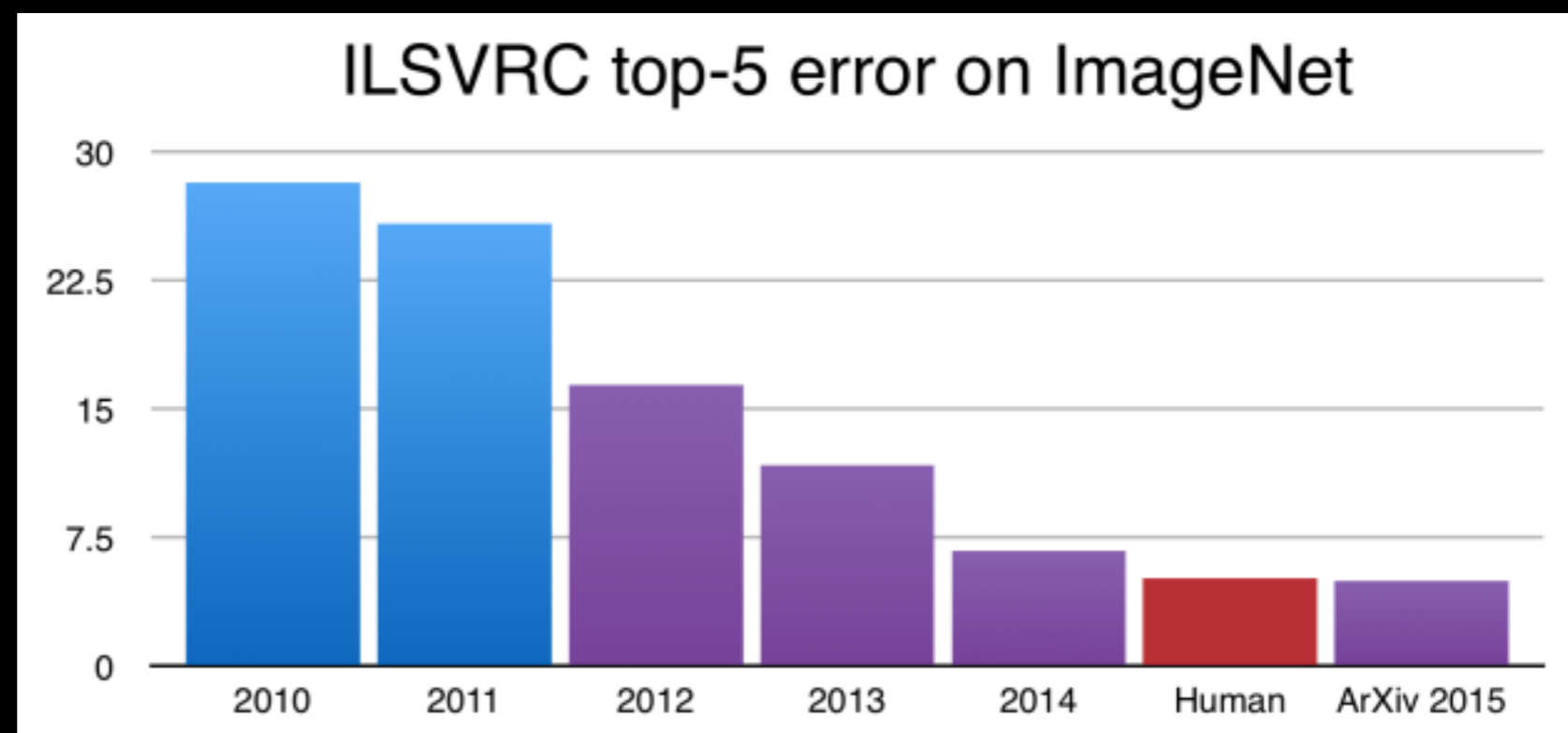
2005: PASCAL VOC challenge: 20.000 images & 20 image classes.

2010: ImageNet competition starts. 1.460.000 images & 1000 image classes.

2012: AlexNet (Alex, Ilya, & Geoffrey) show GPUs are superior for CNNs.

2015: Computers beats humans - in this narrow competition!

[ImageNet Paper](#)



History of the Competition

To identify what was in the images, various teams used Machine Learning (ML).

2005: PASCAL VOC challenge: 20.000 images & 20 image classes.

2010: ImageNet competition starts. 1.460.000 images & 1000 image classes.

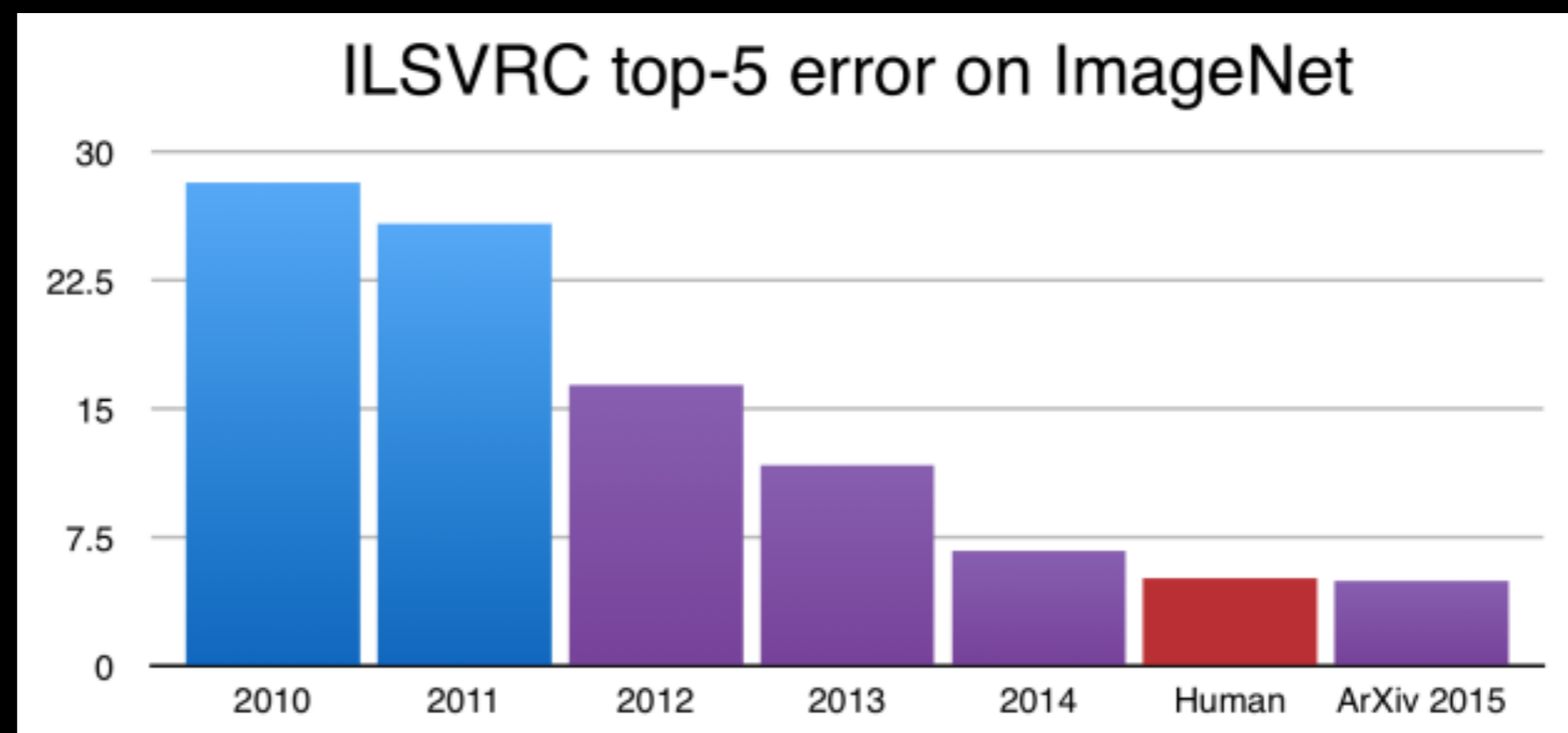
2012: AlexNet (Alex, Ilya, & Geoffrey) show GPUs are superior for CNNs.

2015: Computers beats humans - in this narrow competition!

2017: 29 of 38 competing teams got less than 5% wrong

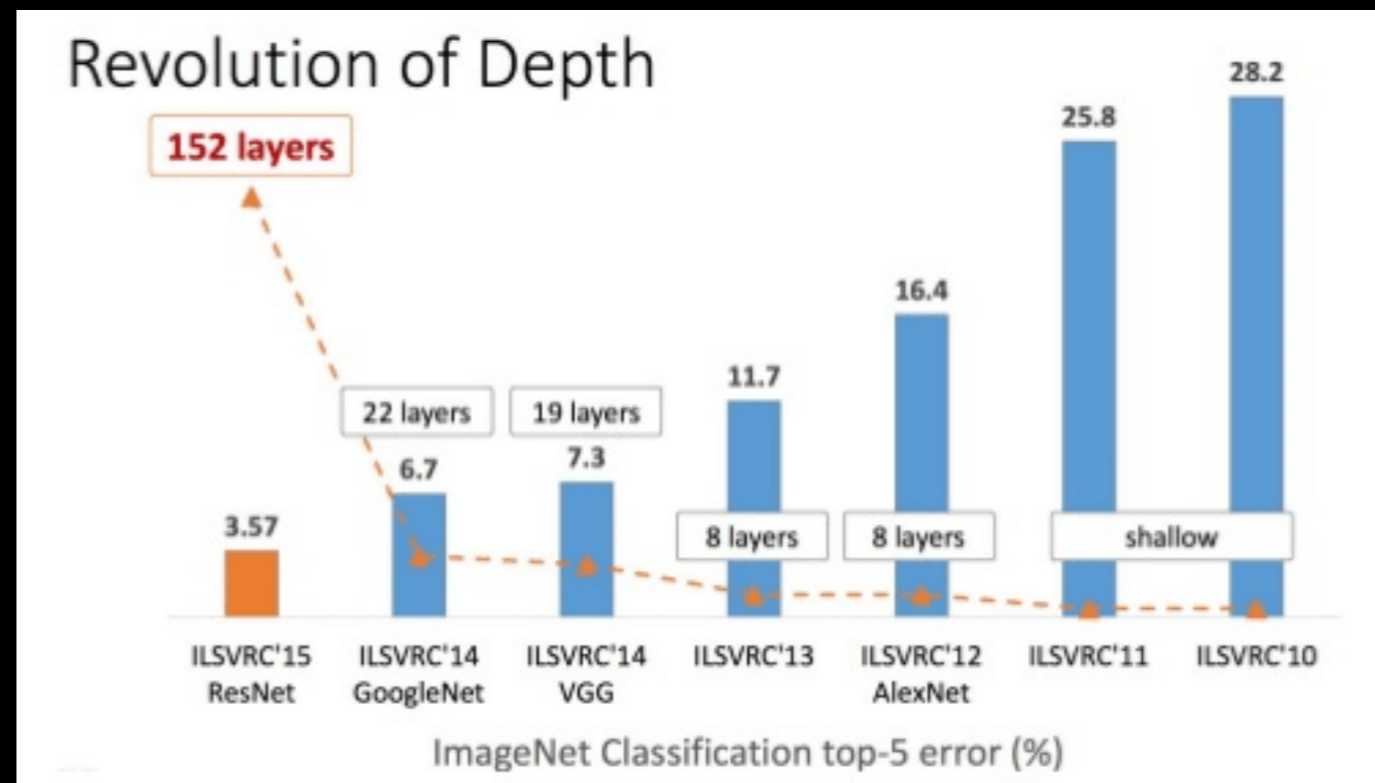
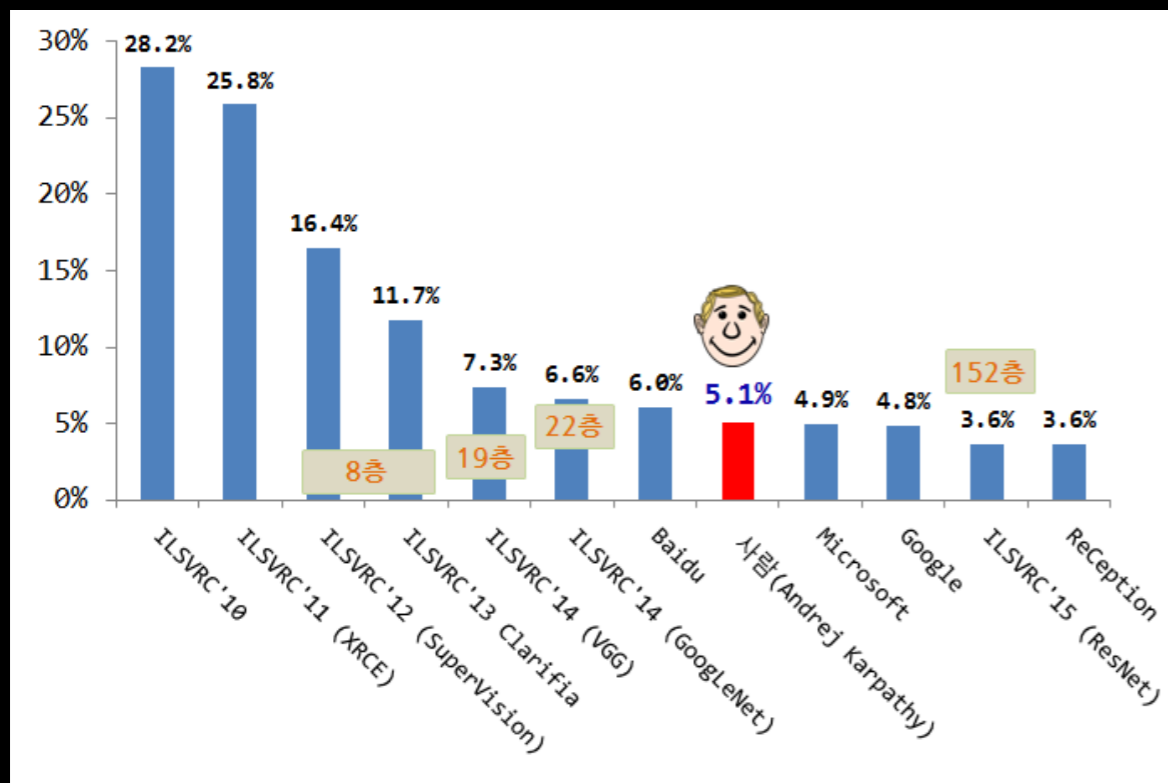
2018+ Challenge changes to describing 3D images in natural language!

[ImageNet Paper](#)



Revolution in Depth

Part of the development was in the realisation, that **deep networks were needed** in order to raise the output in abstraction level.

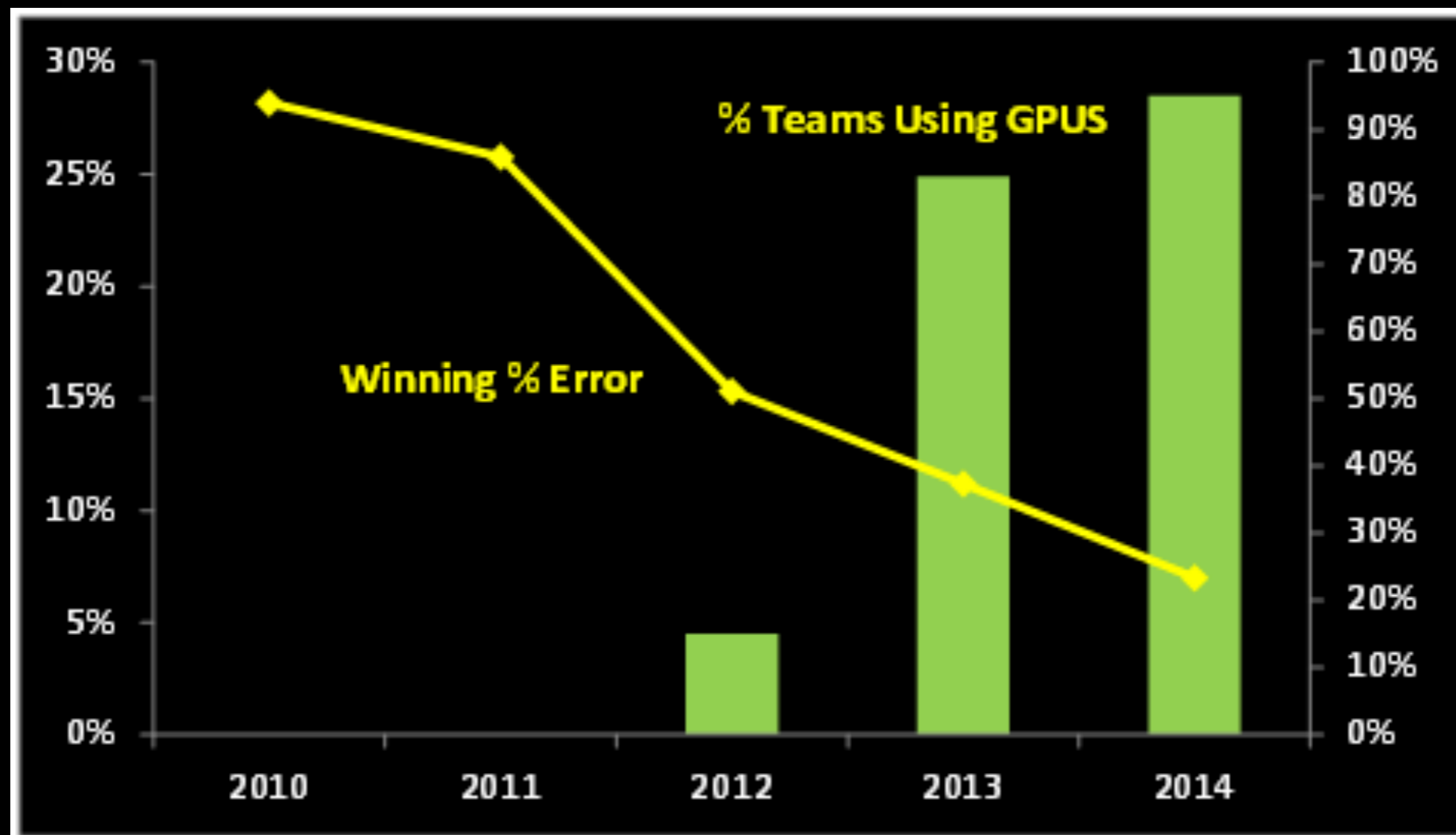


While the 2012 breakthrough "combined pieces that were all there before", the dramatic quantitative improvement marked the start of an industry-wide artificial intelligence boom. [The Economist]

ILSVRC: ImageNet Large Scale Visual Recognition Challenge

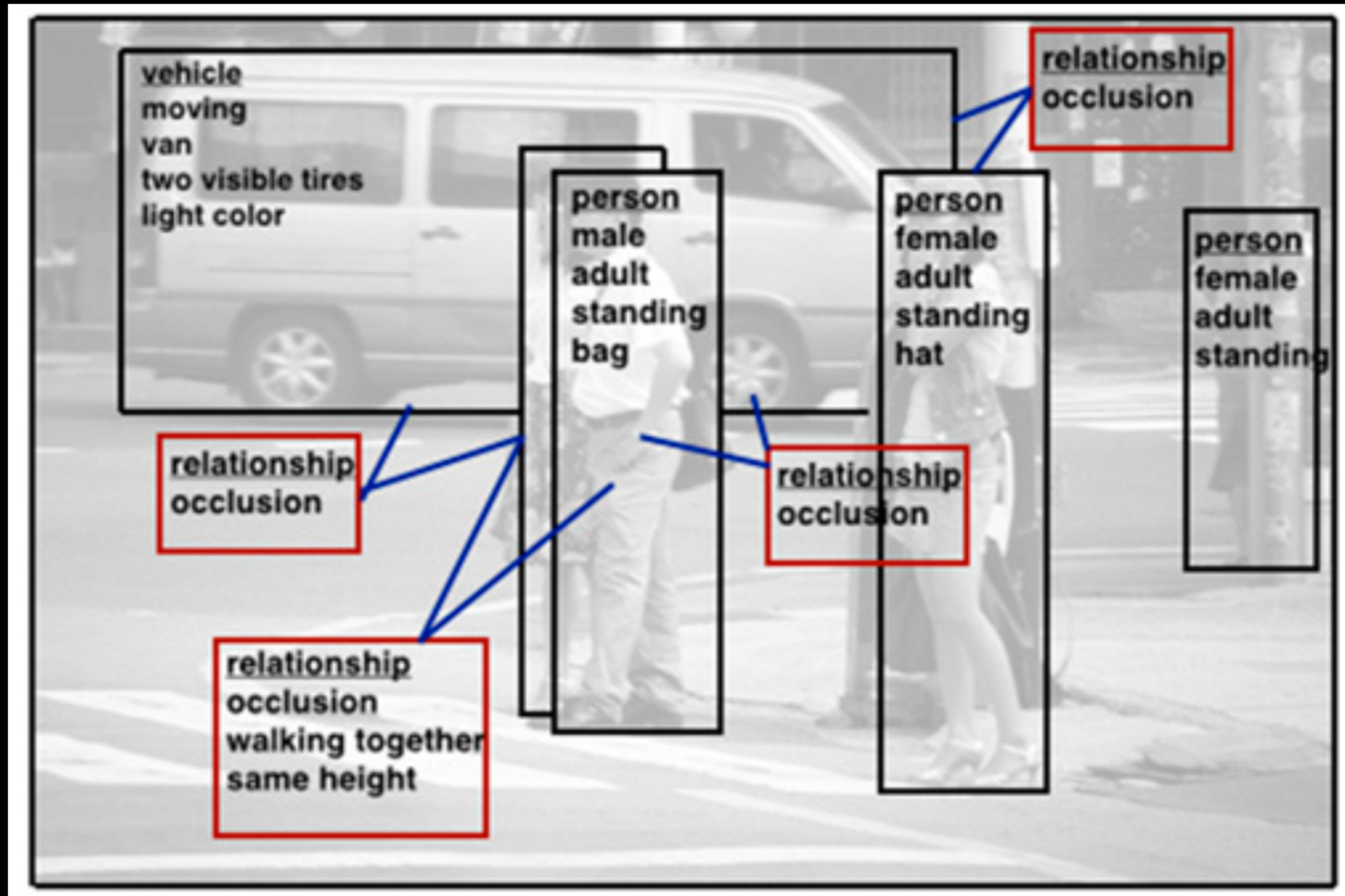
Image ML (CNN) requires GPUs

We have stressed this point several times - it is a task fitted for GPUs not CPUs.



Visual Turing Test

To some extent, the recognition of things in images is a Visual Turing Test...



Raising the abstraction level

Having “solved” the problem of normal 2D images, the abstraction level is raised a level.

In collaboration, Google and Stanford University is aiming at the next steps.

“I consider the pixel data in images and video to be the dark matter of the Internet,”
[Fei-Fei Li, director of Stanford Artificial Intelligence Laboratory]

New York Times,
17. nov. 2014



Human: “Elephants of mixed ages standing in a muddy landscape.”

Computer model: “A herd of elephants walking across a dry grass field.”