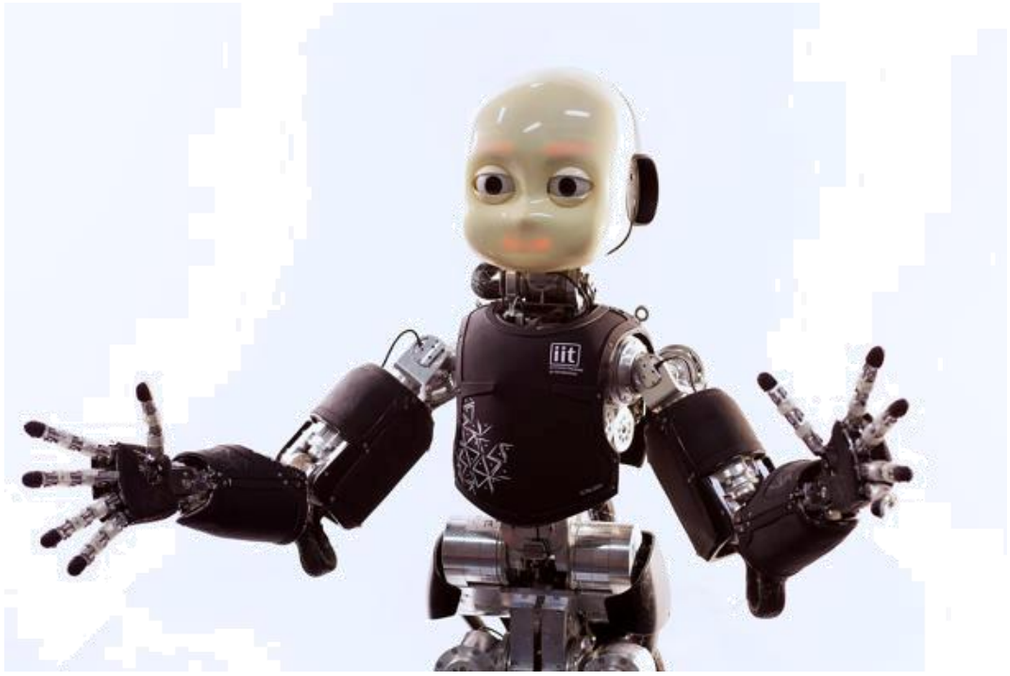
ISTITUTO ITALIANO
DI TECNOLOGIA

LCSL

brains, minds, & machines
summer course 2015 • woods hole, MA



© RobotCub Consortium. All rights reservted. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Giulia Pasquale
PhD student

IIT, iCub Facility –
University of Genoa, DIBRIS –
Laboratory for Computational and Statistical Learning –

ISTITUTO ITALIANO
DI TECNOLOGIA

brains, minds, & machines
summer course 2015 • woods hole, MA

Picture of Lorenzo Natale removed due to copyright restrictions. Please see the video.

Picture of Lorenzo Rosasco removed due to copyright restrictions. Please see the video.

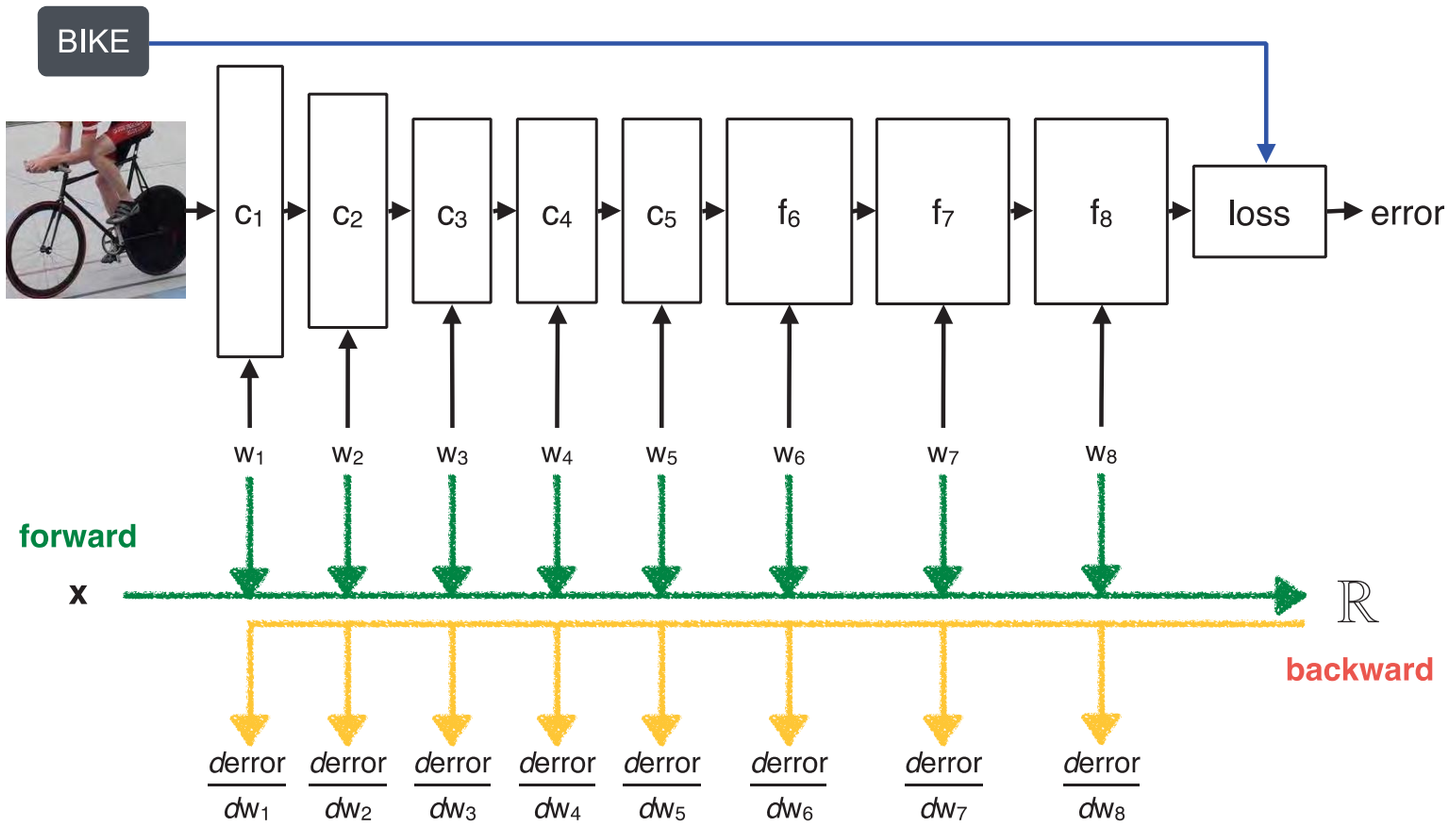
Supervisors and collaborators

Picture of Carlo Ciliberto removed due to copyright restrictions. Please see the video.

Picture of Francesca Odone removed due to copyright restrictions. Please see the video.

Deep Learning Breakthrough in Computer Vision

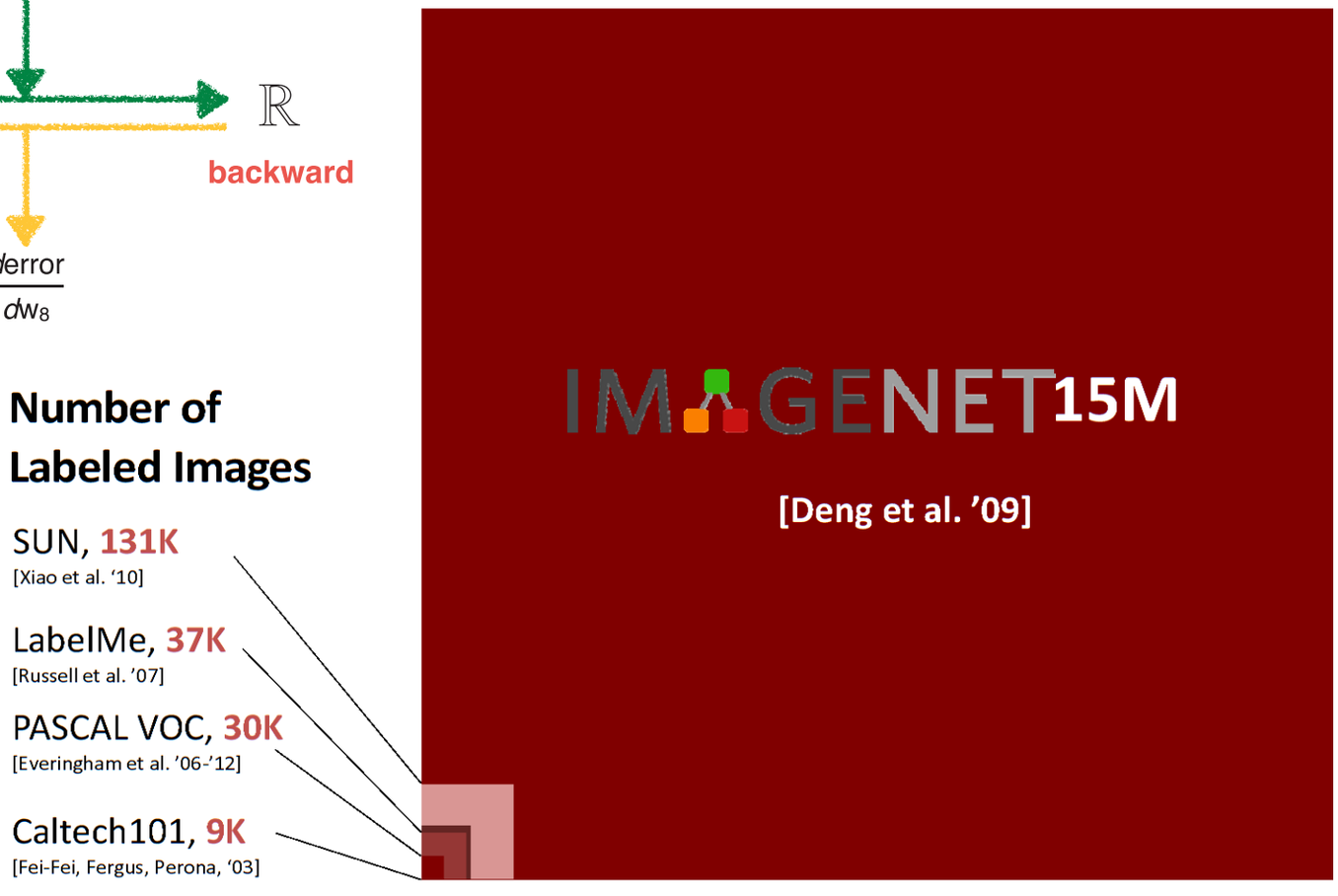
DEEP NETWORKS



Credits: A. Vedaldi

© Oxford Visual Geometry Group. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

BIG DATASETS



Credits: Fei-Fei Li

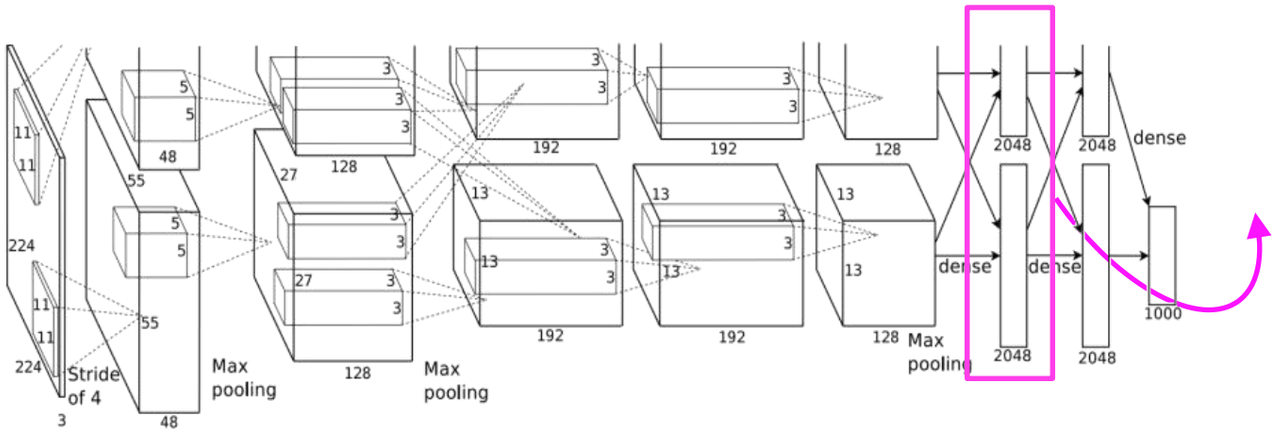
© Source Unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Deep Learning Breakthrough in Computer Vision

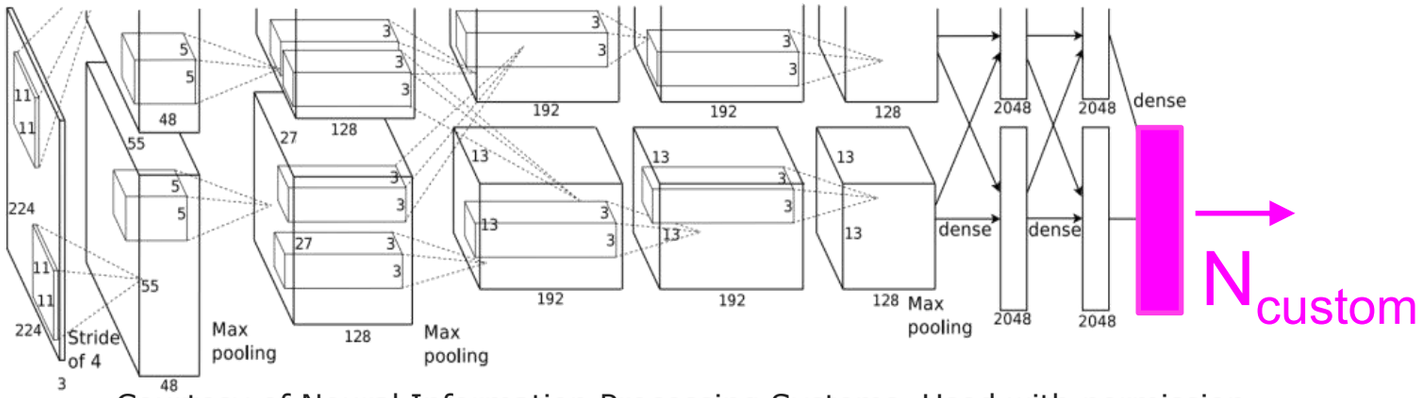
Figure removed due to copyright restrictions. Please see the video.
Source: Figures 9, 11 & 12 from Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang et al. "Imagenet large scale visual recognition challenge." International Journal of Computer Vision 115, no. 3 (2015): 211-252.

Deep Learning Breakthrough in Computer Vision

IMAGENET PRE-TRAINING

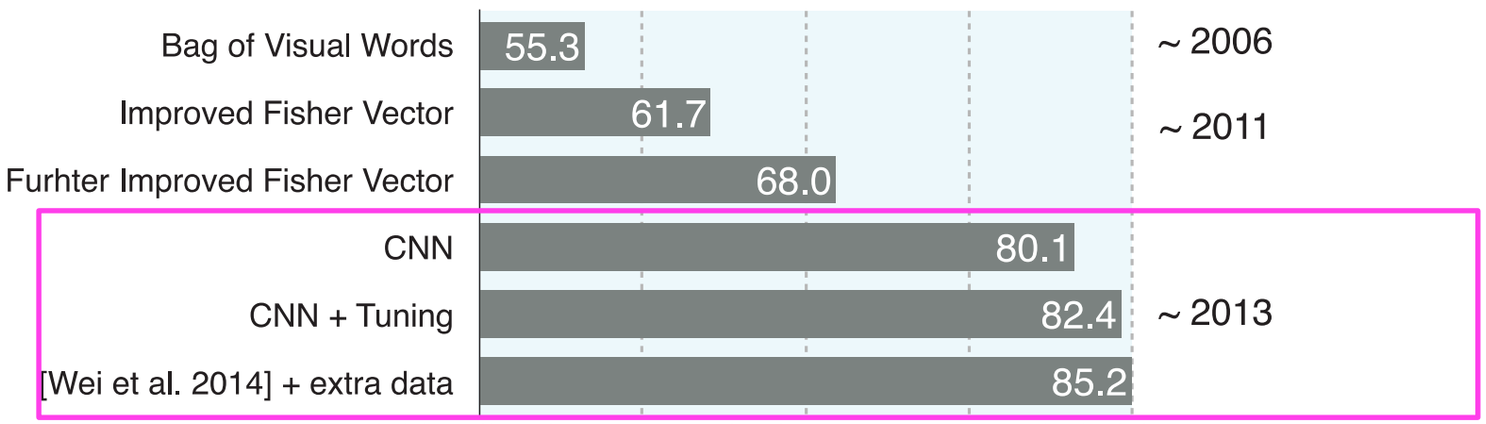


Krizhevsky et al (2012)



Courtesy of Neural Information Processing Systems. Used with permission. Source: Krizhevsky, Alex, IlyaSutskever, and Geoffrey E. Hinton. "Imagene classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.

PASCAL VOC 2007



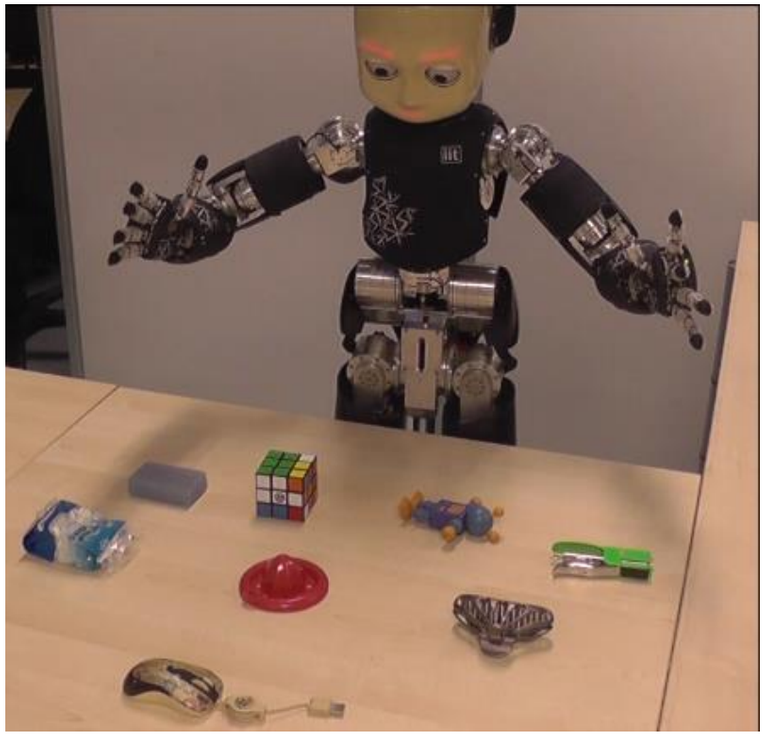
© Andrea Vedaldi. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Credits: A. Vedaldi

Meanwhile, in Robotics…

Image of a baby removed due to copyright restrictions. Please see the video.

Meanwhile, in Robotics…



© RobotCub Consortium. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

© AUVSI. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

© DARPA. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

© DARPA. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

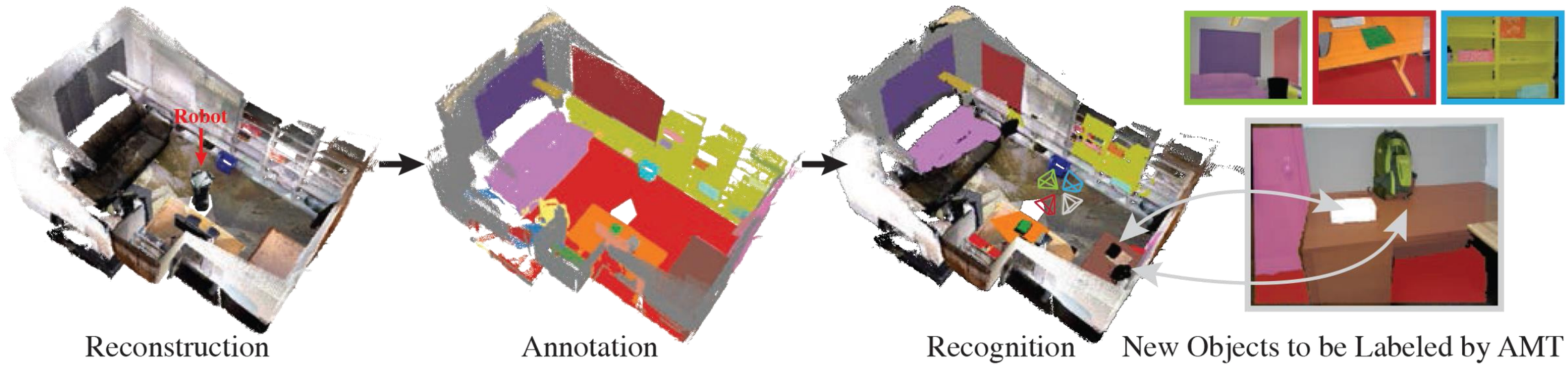
Meanwhile, in Robotics…

TELE-OPERATION

Image removed due to copyright restrictions.
Please see the video.

Image removed due to copyright restrictions.
Please see the video.

3D MAPPING & STRONG SUPERVISION



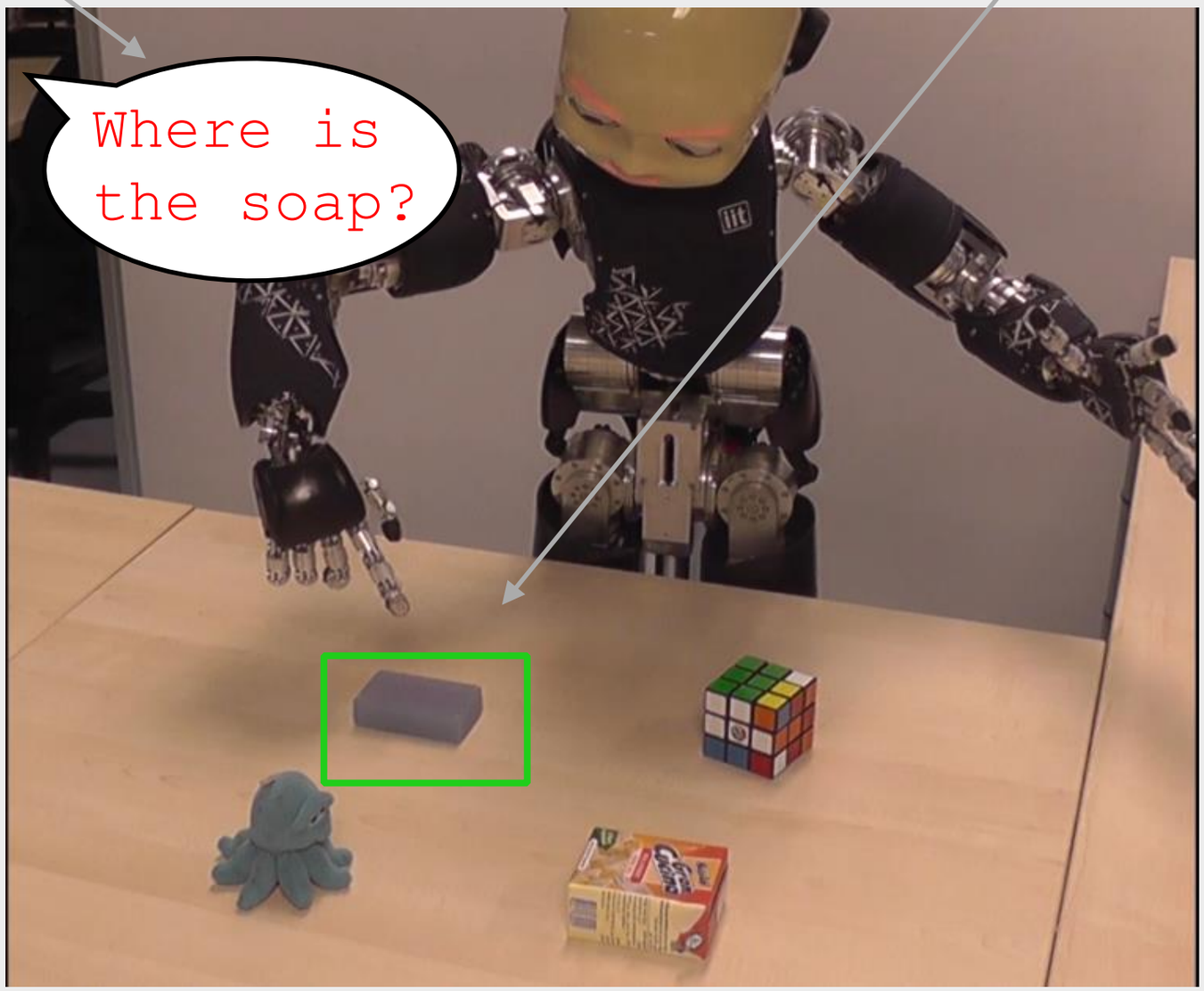
Courtesy of Shuran Song, Linguang Zhang and Jianxiong Xiao. License CC BY.

Song et al (2015), arXiv: 1507.02703

Setting:
Interactive Object Learning

Verbal instructions of a "teacher"

Robot's attention (motion, color-based segmentation)

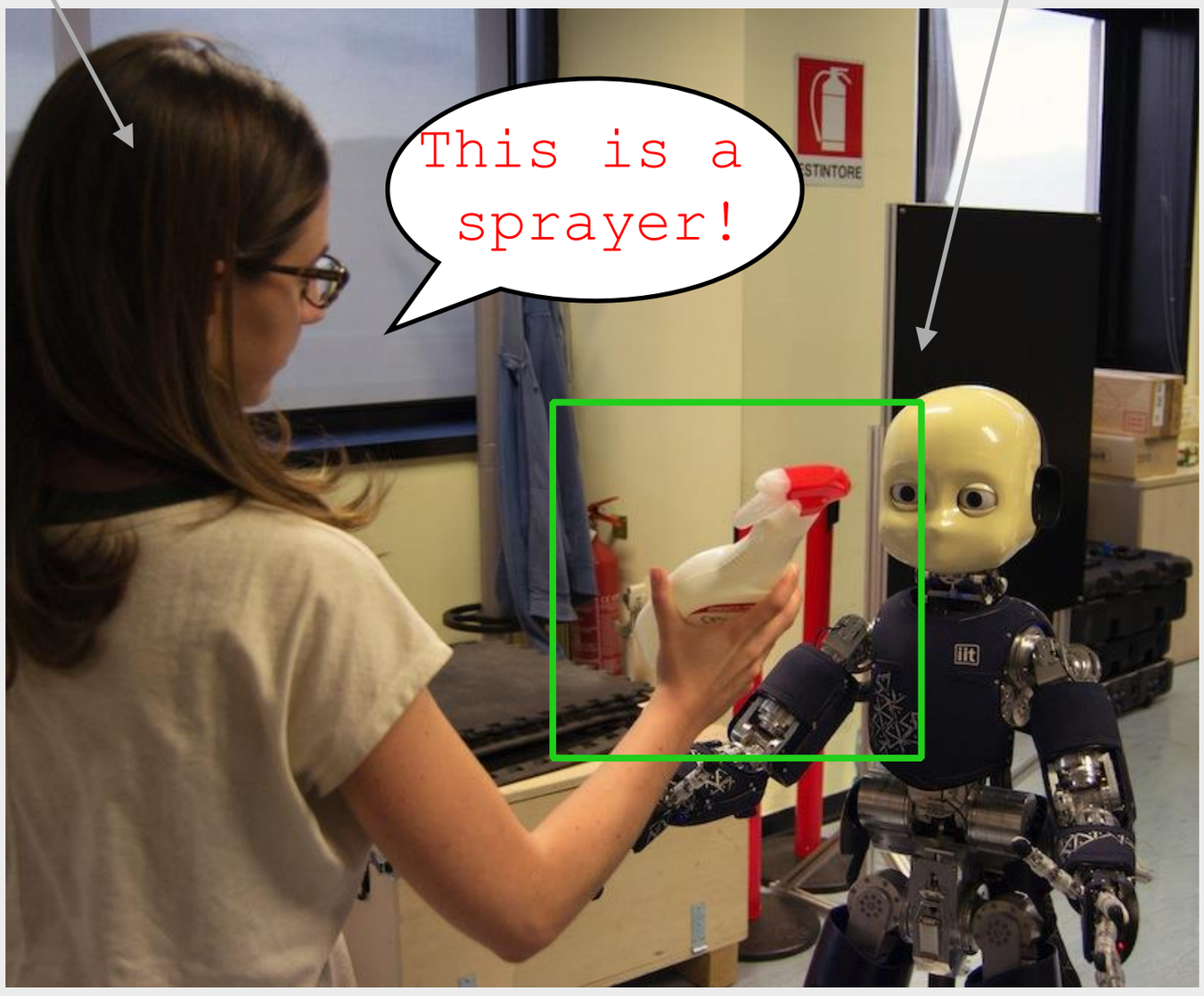


© RobotCub Consortium. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Setting:
On the fly Recognition

Verbal instructions of a "teacher"

Robot's attention (motion)



Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. Used with permission.

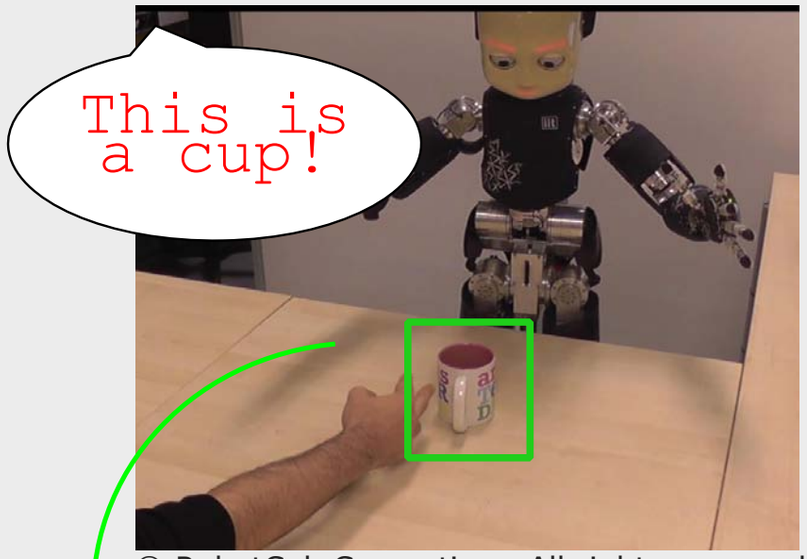
Applications:
Interactive Object Learning & On the fly Recognition

Verbal Supervision

Segmentation

Representation Extraction

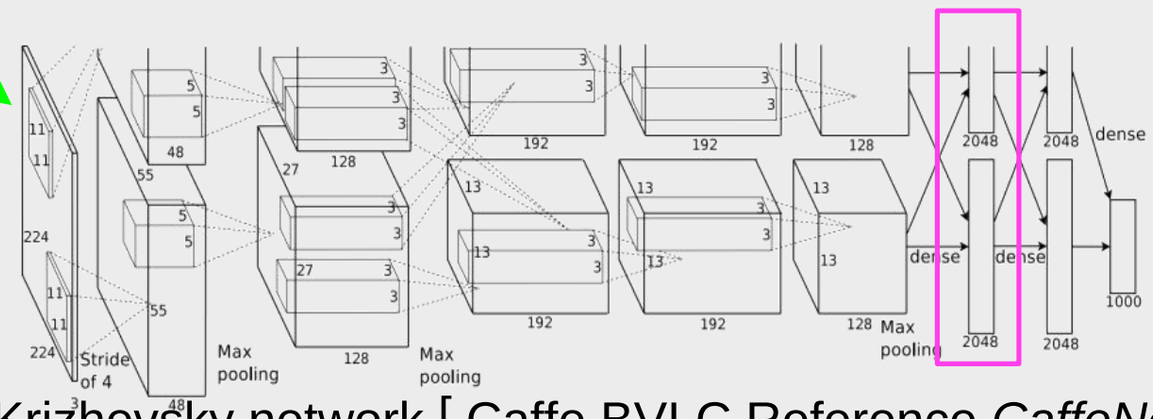
Linear Classifier



Motion, Color & Luminance

© RobotCub Consortium. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

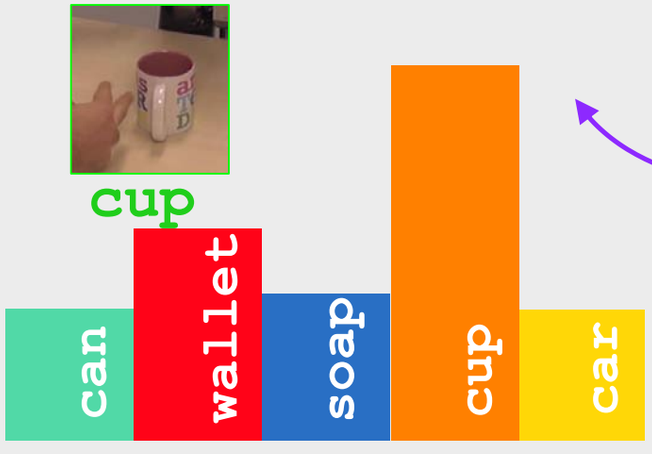
Deep Convolutional Network



Krizhevsky network [Caffe BVLC Reference *CaffeNet*]

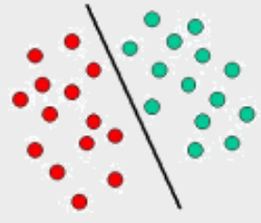
code

Courtesy of Neural Information Processing Systems. Used with permission.
Source: Krizhevsky, Alex, IlyaSutskever, and Geoffrey E. Hinton."Imagene classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.



scores

Linear Classifier



RLS [GURLS]

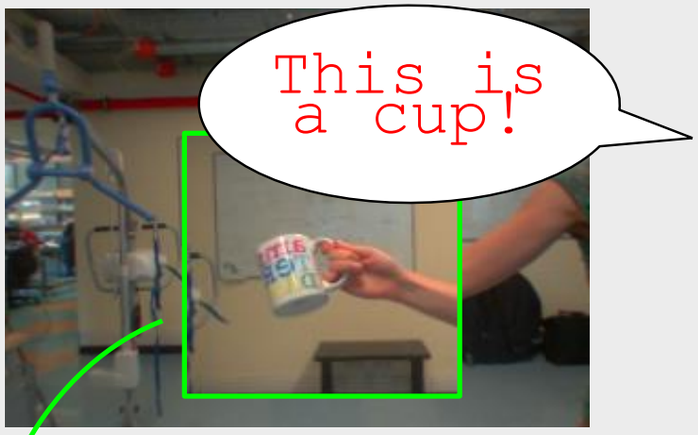
Applications:
Interactive Object Learning & On the fly Recognition

Verbal Supervision

Segmentation

Representation Extraction

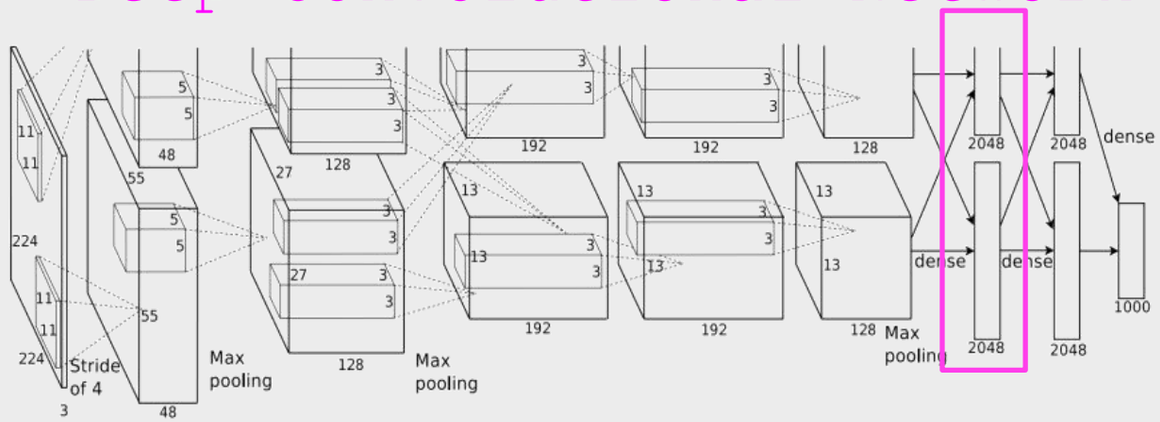
Linear Classifier



Motion, Color & Luminance

Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco, Lorenzo Natale. Used with permission.
Source: Pasquale, Giulia, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco, Lorenzo Natale."Teaching iCubto recognize objects using deep Convolutional Neural Networks." In MLIS@ICML, pp. 21-25. 2015.

Deep Convolutional Network

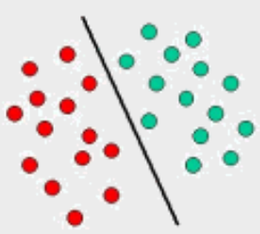


Krizhevsky network [Caffe BVLC Reference *CaffeNet*]

code

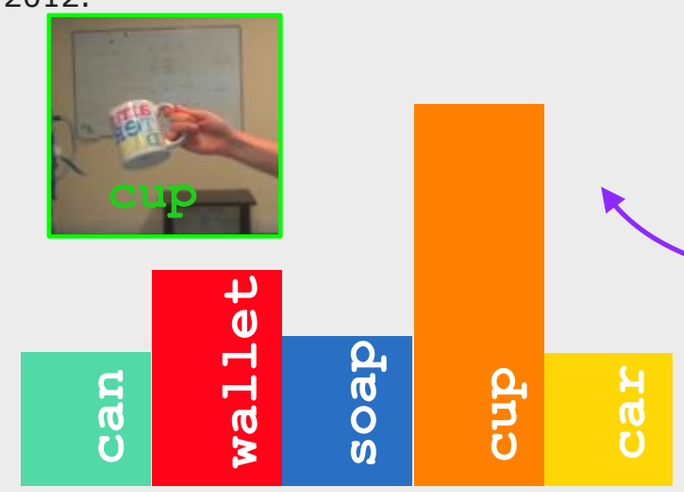
Courtesy of Neural Information Processing Systems. Used with permission.
Source: Krizhevsky, Alex, IlyaSutskever, and Geoffrey E. Hinton. "Imagene classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097-1105. 2012.

Linear Classifier



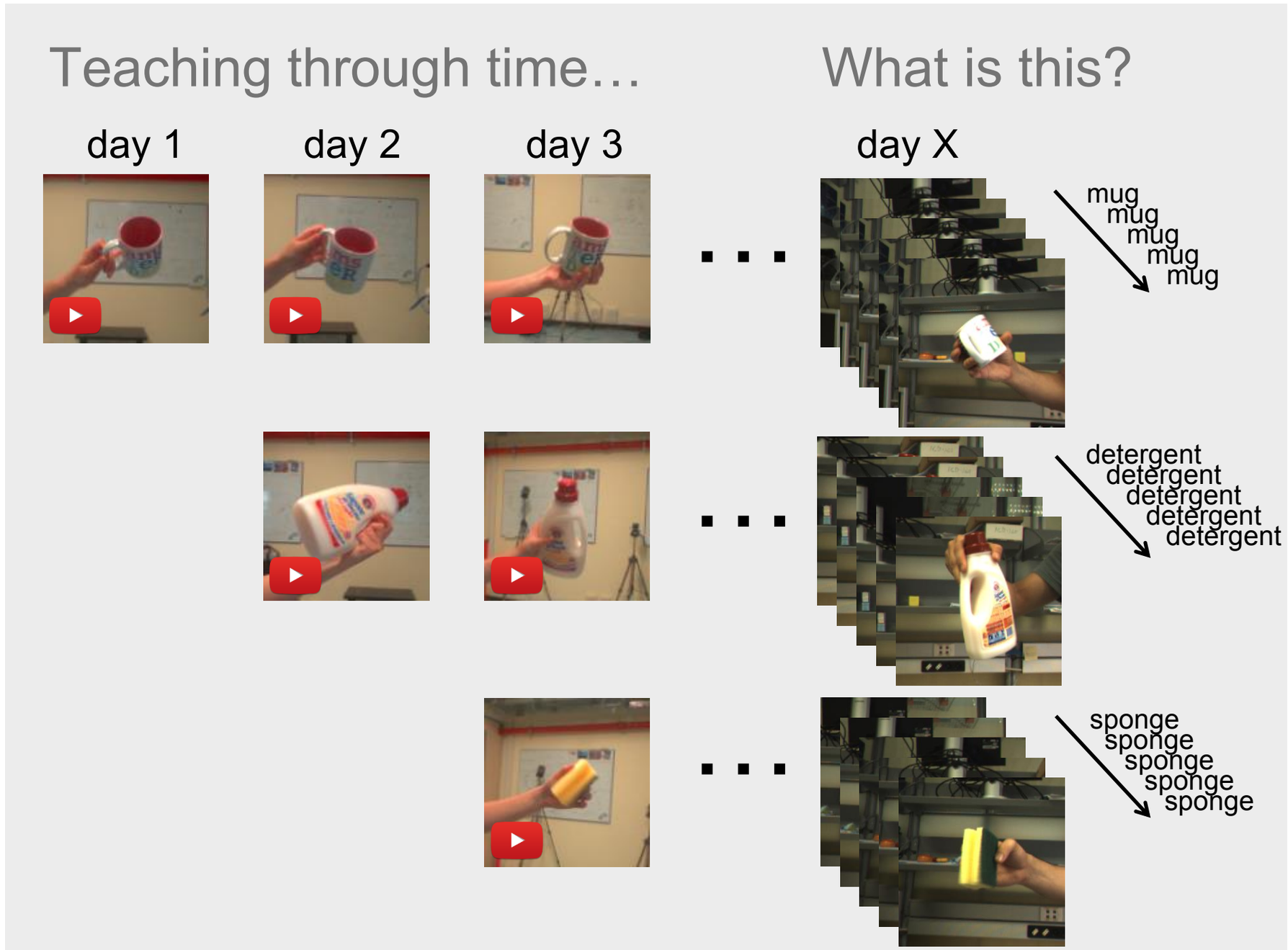
RLS [GURLS]

scores



An ideal robotic visual recognition system

- ✓ Self-supervised
- ✓ Reliable
- ✓ Exploits contextual information
- ✓ Learns incrementally



© Source Unknown. All rights reserved. This content is excluded from ourCreative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

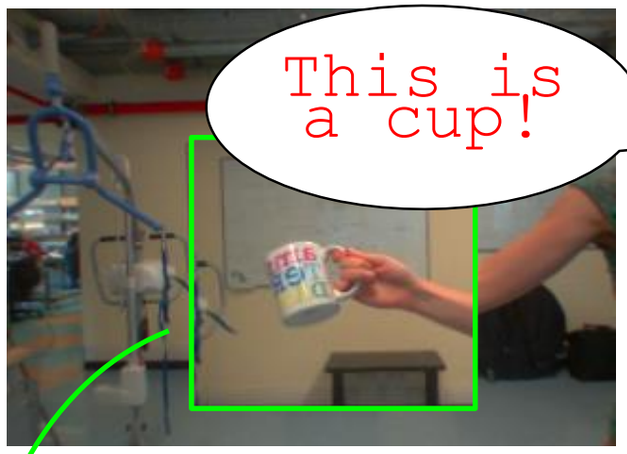
Application:
On the fly Recognition

? Self-supervised

? Reliable

? Exploits contextual information

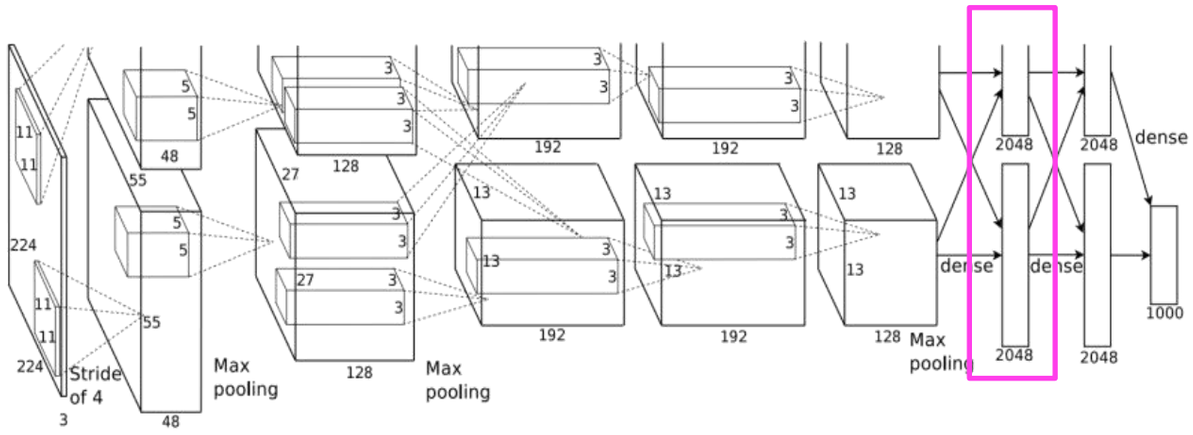
? Learns incrementally



Motion,
Color & Luminance

Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco, Lorenzo Natale. Used with permission.
Source: Pasquale, Giulia, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco, Lorenzo Natale. "Teaching iCubto recognize objects using deep Convolutional Neural Networks." In MLIS@ICML, pp. 21-25. 2015.

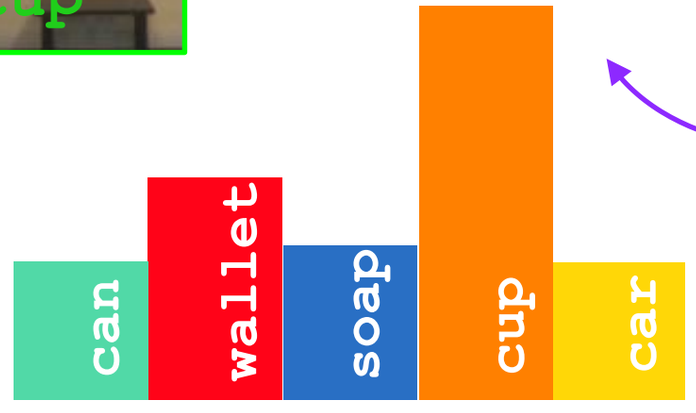
Deep Convolutional Network



Krizhevsky network [Caffe BVLC Reference *CaffeNet*]

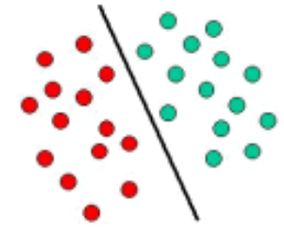
code

© Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.



scores

Linear Classifier



RLS [GURLS]

iCubWorld28 Dataset

Overview

2014: ‘‘Household’’

7 categories
4 objects/category
→ 28 objects

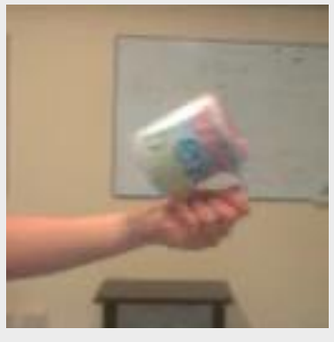
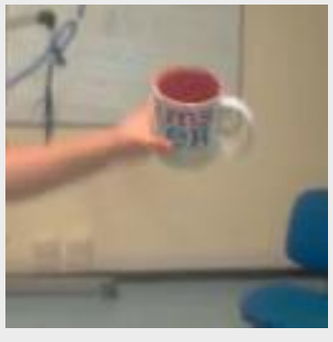
4 acquisitions

Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. Used with permission.

Benchmarking deep Conv Nets for Real-world Object Recognition: How many Objects can iCub Learn?
arXiv: 1504.03154,

iCubWorld28 Dataset
Examples of Acquired Videos

2014: ‘‘Household’’

	day1	day2	day3	day4
TRAIN				
TEST				

© Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale.
All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/.

Benchmarking deep Conv Nets for Real-world Object Recognition: How many Objects can iCub Learn?
arXiv: 1504.03154,

iCubWorld28 Dataset

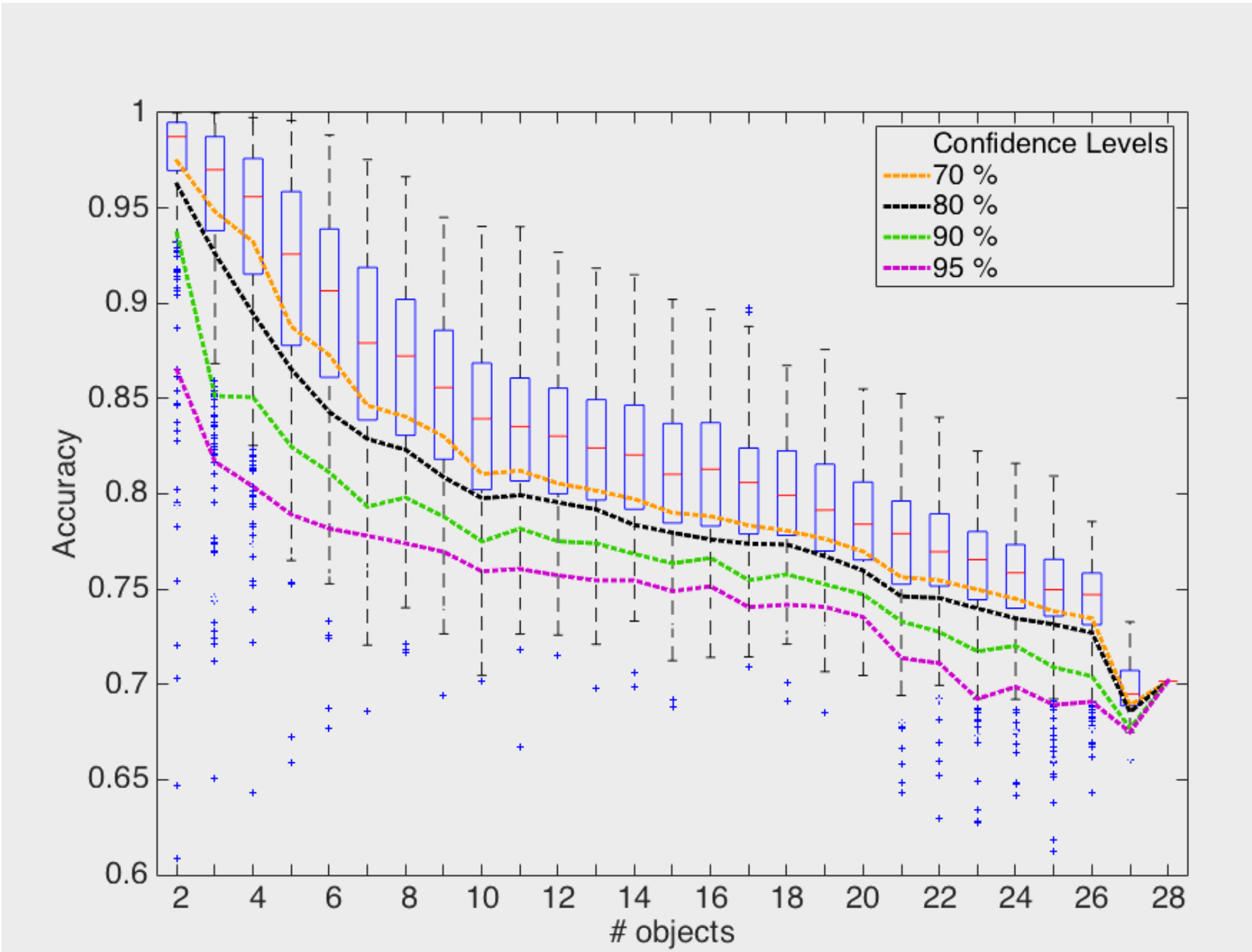
Object Identification “Data Sheet”

? Self-supervised

? Reliable

? Exploits contextual information

? Learns incrementally



Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. Used with permission.

Benchmarking deep Conv Nets for Real-world Object Recognition: How many Objects can iCub Learn?
arXiv: 1504.03154,

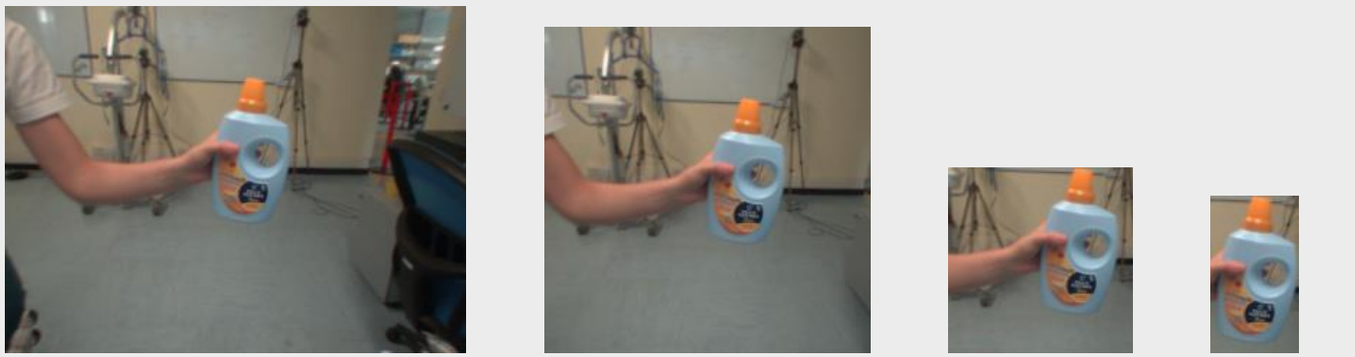
iCubWorld28 Dataset
Clutter and Scale

? Self-supervised

? Reliable

? Exploits contextual information

? Learns incrementally



		TEST Accuracy (%)			
		Image	Crop1	Crop2	Manual
TRAIN	Image	50.6	48.8	36.3	20.6
	Crop1	50.3	62.2	57.7	24.9
	Crop2	30.1	50.8	73.9	28.7
	Manual	6.8	8.9	12.2	**81.7**

Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. Used with permission.

Benchmarking deep Conv Nets for Real-world Object Recognition: How many Objects can iCub Learn?
arXiv: 1504.03154,

iCubWorld28 Dataset

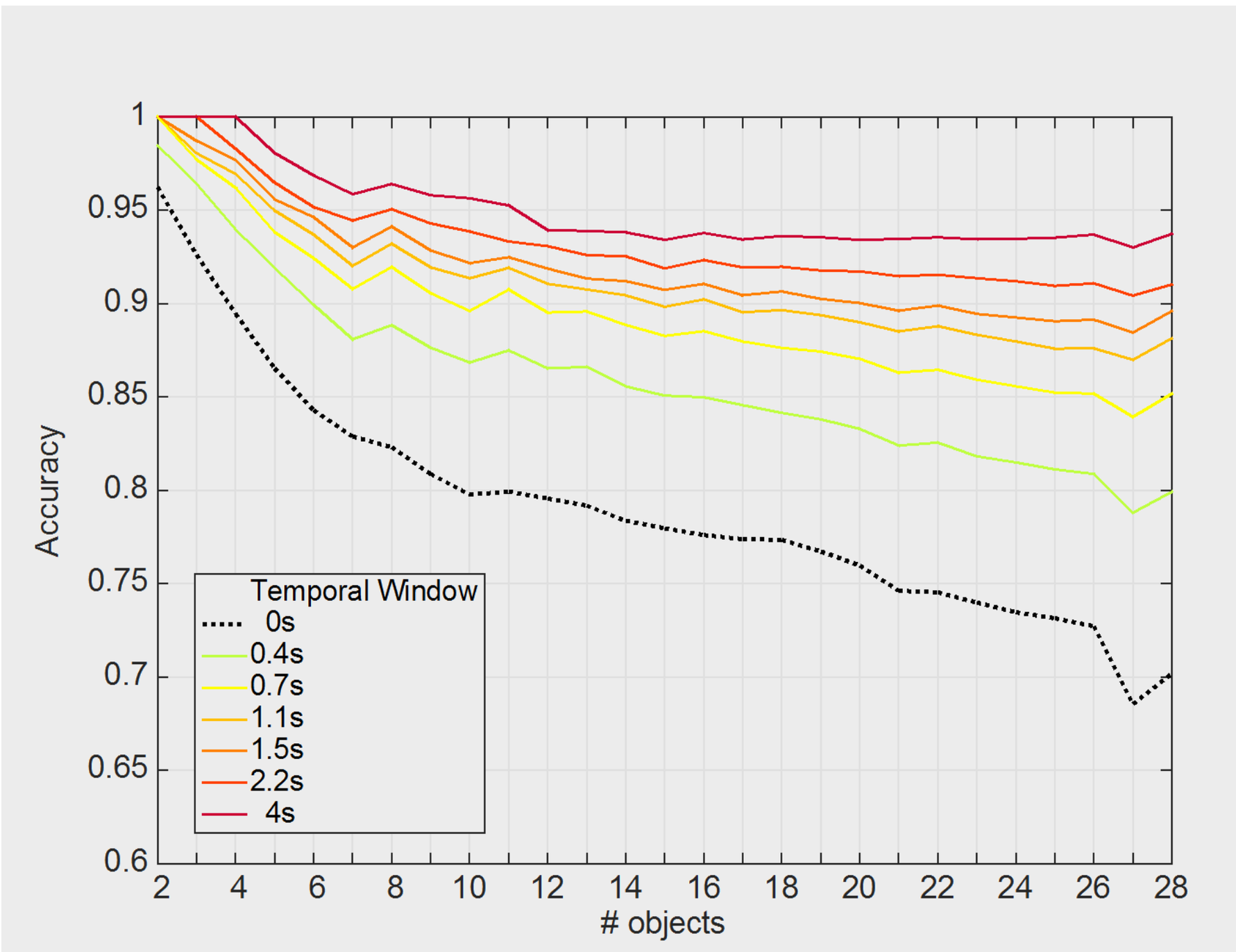
Temporal Contextual Information

? Self-supervised

? Reliable

? Exploits contextual information

? Learns incrementally



Courtesy of Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco and Lorenzo Natale. Used with permission.

Benchmarking deep Conv Nets for Real-world Object Recognition: How many Objects can iCub Learn?
arXiv: 1504.03154,

iCubWorld Datasets
Ongoing Work

2015:
“Kitchen” + “Food” + “Toys” + “Bathroom” + “Daily use” + “Office” + “Tools”

oven glove	bottle	beach:	body cream	sunglasses	mouse	paint brush
squeezer	box	shovel	hair brush	hair clip	organizer	scissors
cup	can	bucket	soap	watch	calculator	scotch
		toy rake				

1. **Object Categorization Dataset**
   - **21 categories**
   - **10 objects/category**
   - **→ 200 objects**
2. **Continuously Expandable in Time**
3. **Tagged by Nuisance Factors**
   - **5 acquisitions divided by nuisance:**
     - ✓ **scale**
     - ✓ **2D rotation**
     - ✓ **3D rotation**
     - ✓ **translation**
     - ✓ **mixed**
4. **Depth information available (left+right cameras)**

iCubWorld Datasets

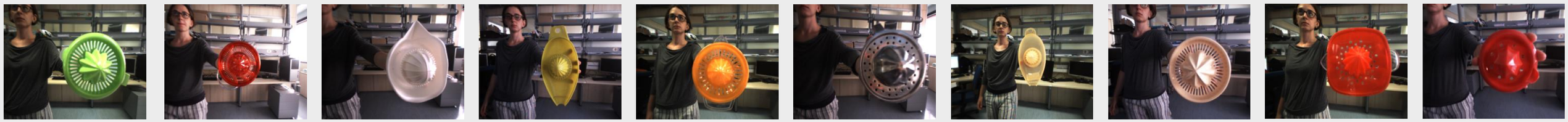
Disparity-driven segmentation

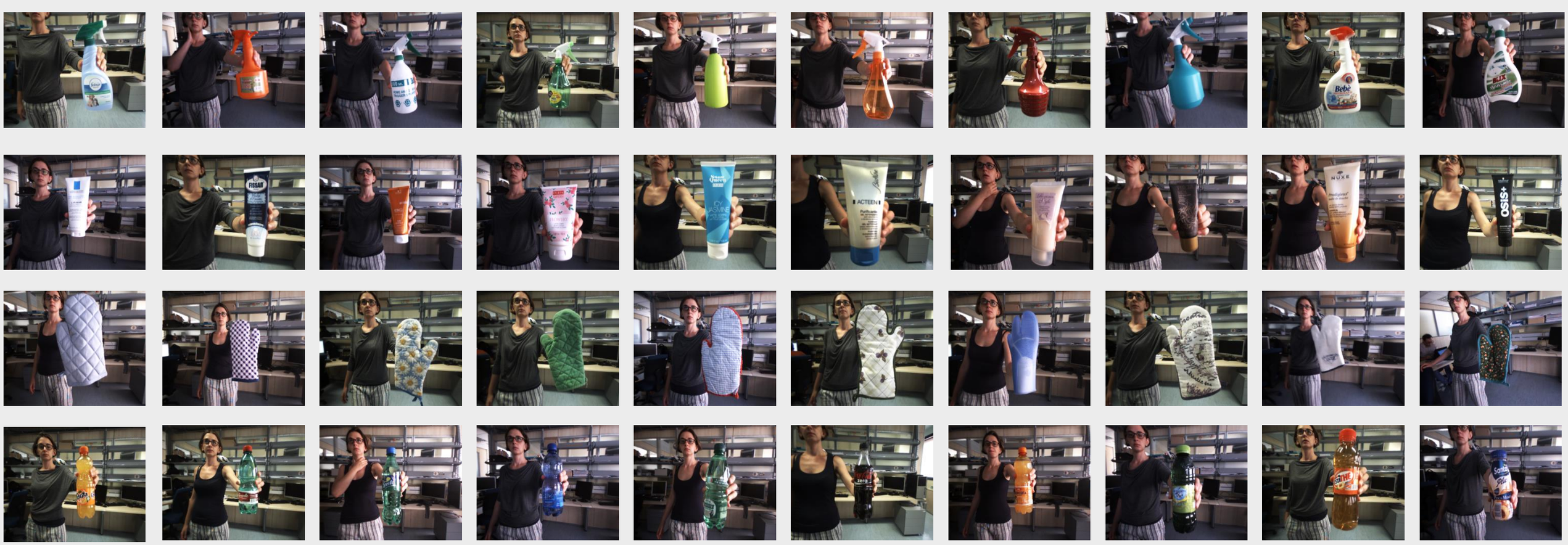
Courtesy of Giulia Pasquale, Tanis Mar, Carlo Ciliberto, Lorenzo Rosasco, and Lorenzo Natale. Used with permission.

iCubWorld Datasets
Ongoing Work

2015:
“Kitchen” + “Food” + “Toys” + “Bathroom” + “Daily use” + “Office” + “Tools”

oven glove	bottle	shovel	body cream	sunglasses	mouse	paint brush
squeezer	box	bucket	hair brush	hair clip	organizer	scissors
cup	can	toy rake	soap	watch	calculator	scotch





iCubWorld Datasets
Ongoing Work

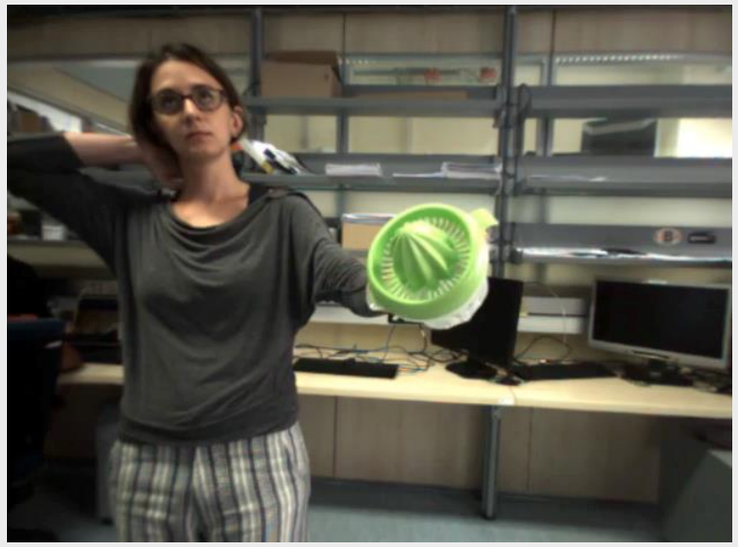
2015:
“Kitchen” + “Food” + “Toys” + “Bathroom” + “Daily use” + “Office” + “Tools”

translation

scale

2D rotation

3D rotation



mixed

Application & Data are available for projects 5.2 & 5.3!!

THANK
You !

MIT OpenCourseWare
https://ocw.mit.edu

Resource: Brains, Minds and Machines Summer Course

Tomaso Poggio and Gabriel Kreiman

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.