

Course Logistics

CS 783: Visual Recognition

Vinay P. Namboodiri
vinaypn@iitk.ac.in

Instructor Details

- Vinay P. Namboodiri
- # RM 406, RM building, CSE
- Office Hours: Monday 10 - 11 am or by appointment
- Preferably email: vinaypn@iitk, Subject: CS783....

Teaching Assistants

- Aman Deep Singh: amands@iitk.ac.in
- Pravendra Singh: psingh@iitk.ac.in
- Saket Jhunjhunwala: saketj@cse.iitk.ac.in
- Samik Some: samiks@iitk.ac.in
- Siddharth Singla: ssid@cse.iitk.ac.in
- Utsav Singh: utsavz@cse.iitk.ac.in

Course Details

- Course Website:
- <https://www.cse.iitk.ac.in/users/vinaypn/teaching/cs783/>
- Piazza link
- <https://piazza.com/iitk.ac.in/secondsemester2019/cs783>

About the Course

- Visual Recognition
- A graduate elective
- Objectives: Obtain different perspectives on understanding visual recognition
 - Problems
 - Approaches
 - Advances

Lectures

- Tuesday, and Thursday
- Tuesday 17:10 -18:25, Thu 17:10-18:25,
- Venue: L19

Grading

- Weightage (Tentative):
- Mid-Sem - 25%
- End-Sem 25%
- Assignments and Quizzes 25% (Programming, Paper Review)
- Project 25% (atleast 2 stages)

Course Outline

- Introduction
- Exact instance retrieval
- Classification
- Detection
- Segmentation
- Weak Supervision
- Active Learning
- Domain Adaptation
- Unsupervised Representation learning
- Vision and Language

Course Outline

- Introduction
- Exact instance retrieval
- Classification
- Detection
- Segmentation
- Weak Supervision
- Active Learning
- Domain Adaptation
- Unsupervised Representation learning
- Vision and Language

Course Outline

- Introduction
- Exact instance retrieval
- Classification
- Detection
- Segmentation
- Weak Supervision
- Active Learning
- Domain Adaptation
- Unsupervised Representation learning
- Vision and Language

Traditional
Feature
Based

Course Outline

- Introduction
- Exact instance retrieval
 - Classification
 - Detection
 - Segmentation
- Weak Supervision
- Active Learning
- Domain Adaptation
- Unsupervised Representation learning
- Vision and Language

Deep
learning
based

Course Outline

- Introduction
- Exact instance retrieval
- Classification
- Detection
- Segmentation

Deep
learning
based

- Weak Supervision
- Active Learning
- Domain Adaptation
- Unsupervised Representation learning
- Vision and Language

Tentative set
of
advanced topics

Course Material

- Lecture slides that will be posted online
- Course will be based mainly on research papers
- Reference books:
- Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville Available online
- Computer Vision: Algorithms and Applications by Richard Szeliski Available online
- Computer Vision: Models, Learning, and Inference by Simon J.D. Prince Available online
- Computer Vision: A Modern Approach by Forsyth and Ponce Indian edition available

Introduction

CS 783: Visual Recognition

Vinay P. Namboodiri
vinaypn@iitk.ac.in

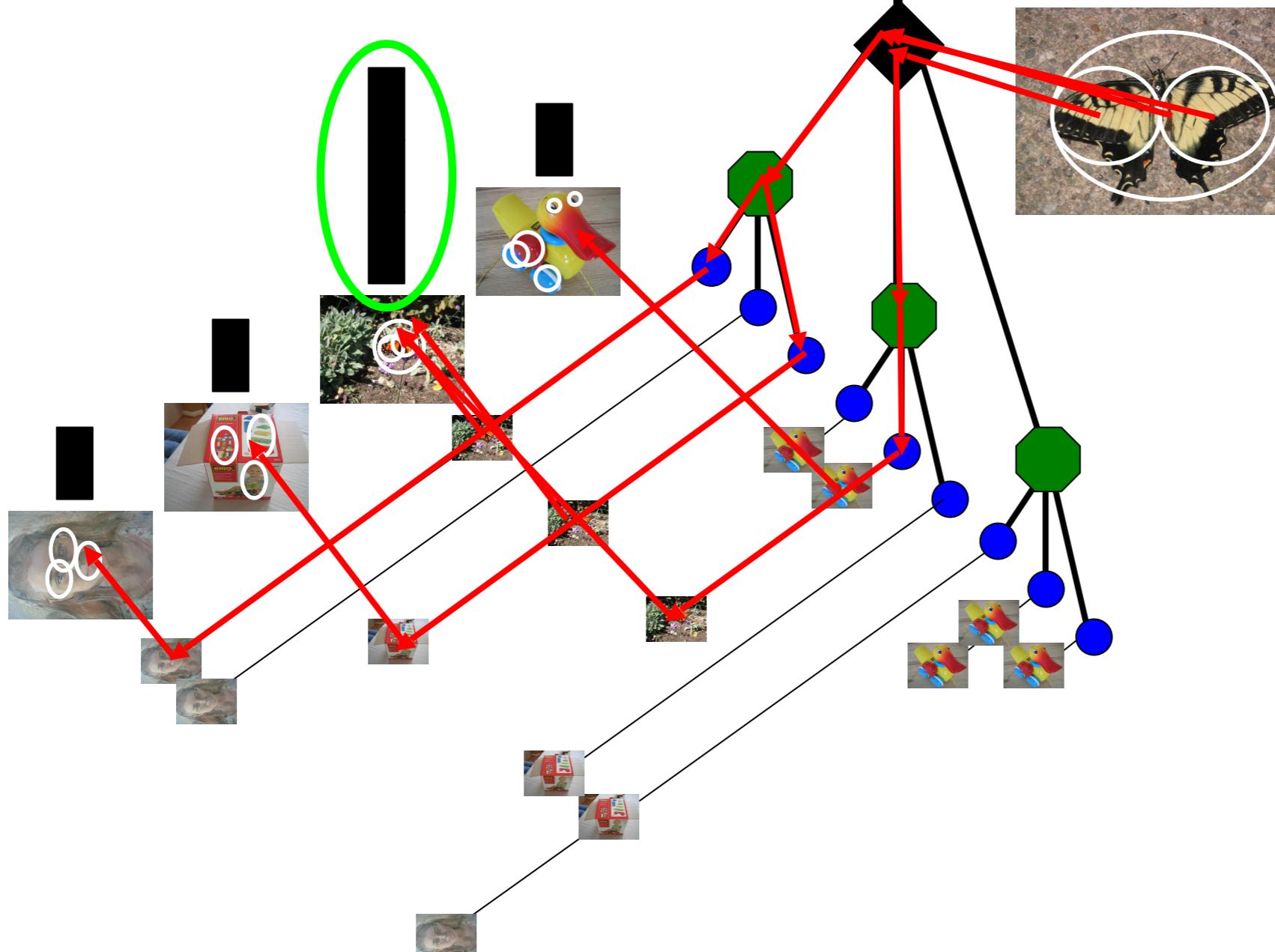
What is Visual Recognition?

What is Visual Recognition

- Definition:[1]:
- It is the ability to perceive an object's physical properties (such as shape, colour and texture) and apply semantic attributes to it (such as identifying the object as an apple).
- This process includes the understanding of its use, previous experience with the object, and how it relates to others
- Regardless of an object's position or illumination, humans possess the ability to effectively identify and label an object.
- Humans are one of the few species that possess the ability of invariant visual object recognition.
- Both "front end" (knowledge/goal driven) and "back end" (sensory driven) processing are required for a species to be able to recognize objects at varying distances, angles, lighting, etc....



Semantic Representation



Instance Recognition

Scalable recognition with a Vocabulary tree

Nister, Stewenius, CVPR 2006



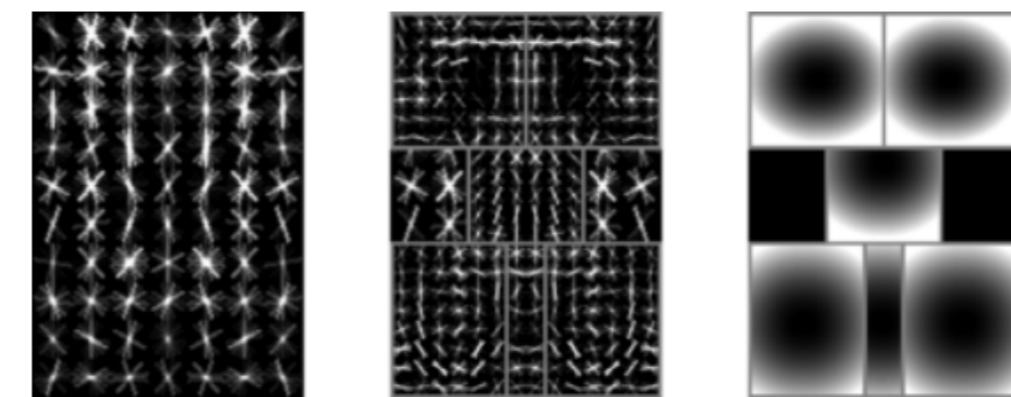
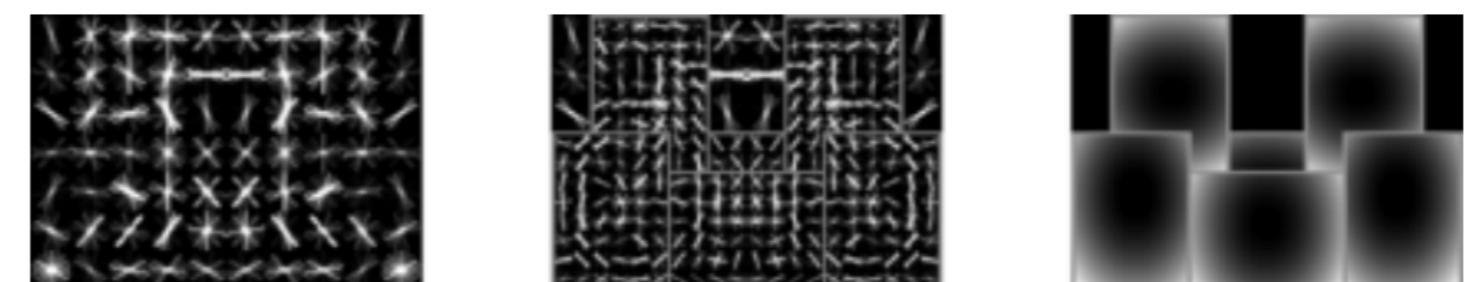
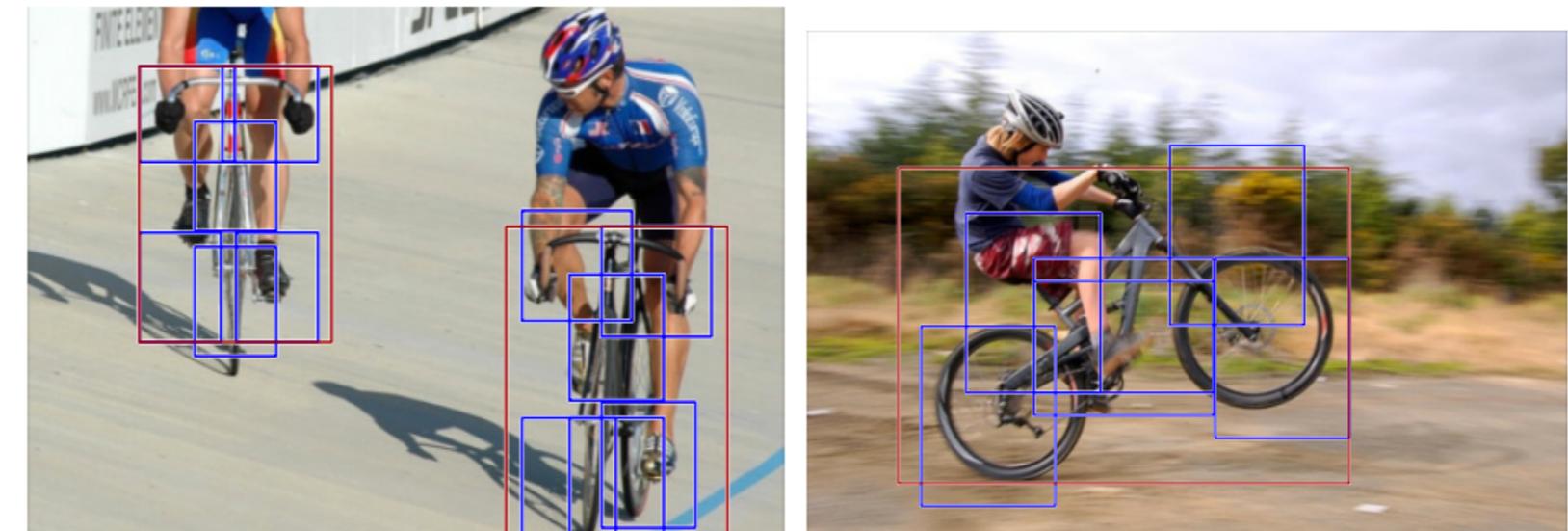
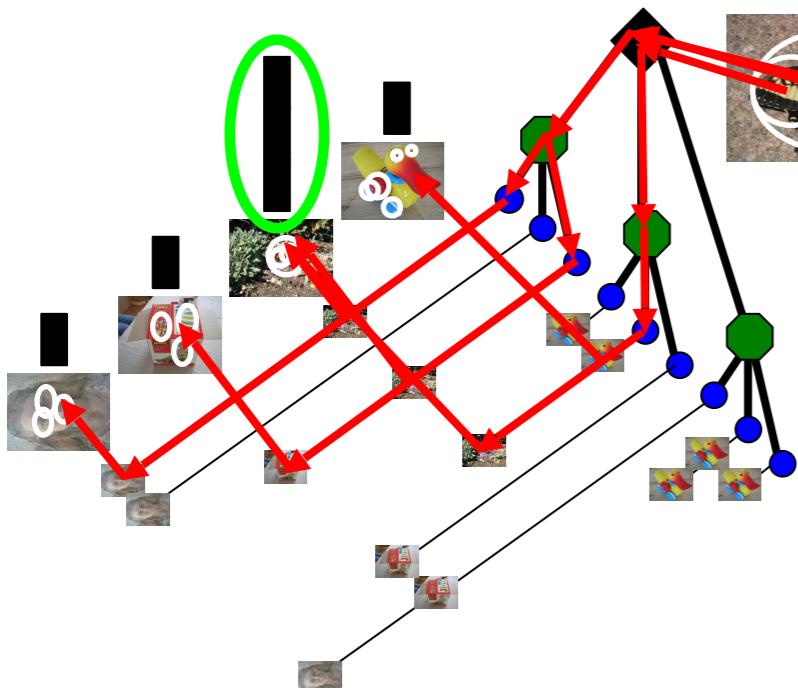
Semantic Representation



Object Classification
ImageNet

Image credit: Karpathy

Semantic Representation

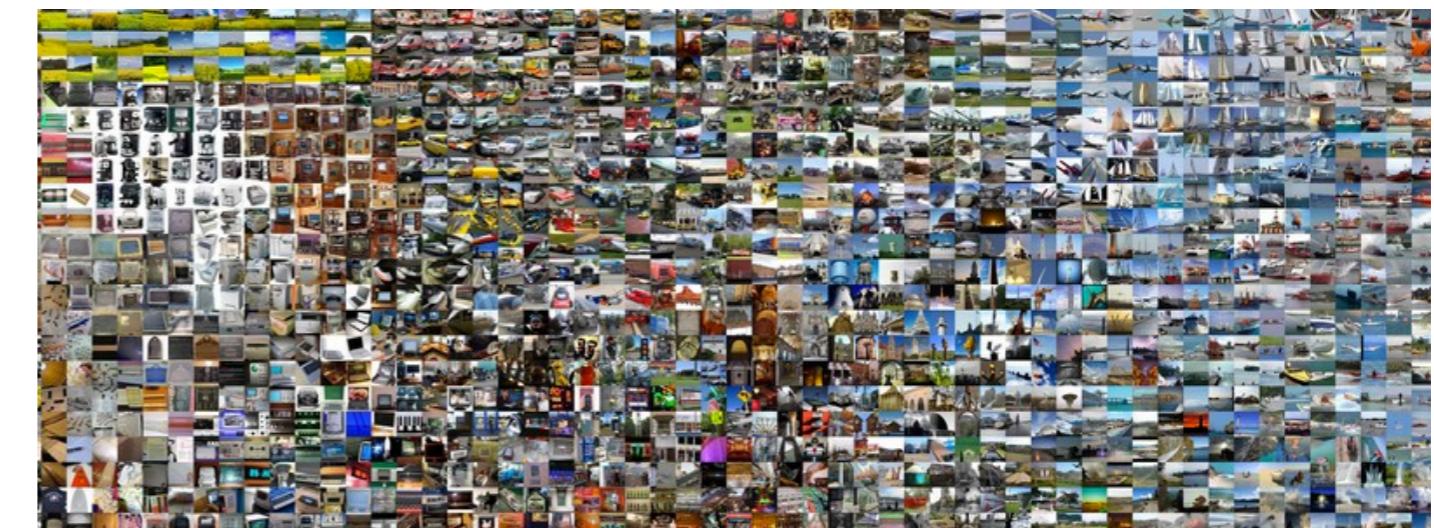
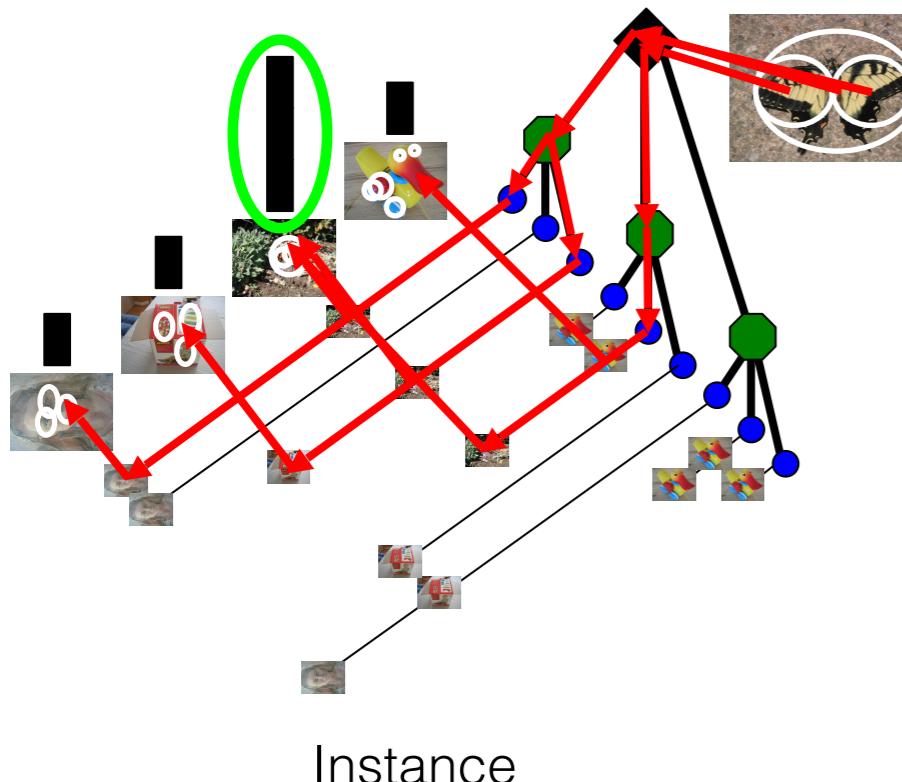


Object Detection

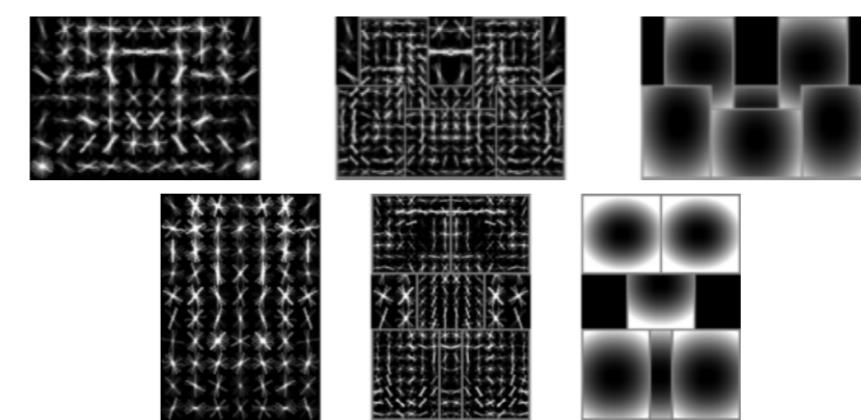
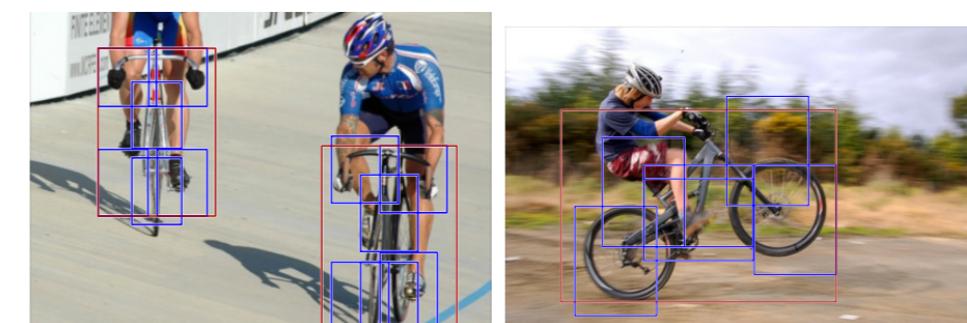
Object Detection with Discriminatively Trained Part Based Models
P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan
PAMI 2010



Semantic Representation



Object



Object Detection

Why understand
Visual Recognition?

Motivation

- Intellectual curiosity
- Algorithms for general visual perception (also enable general machine learning methods)
- Applications

Intellectual Challenge

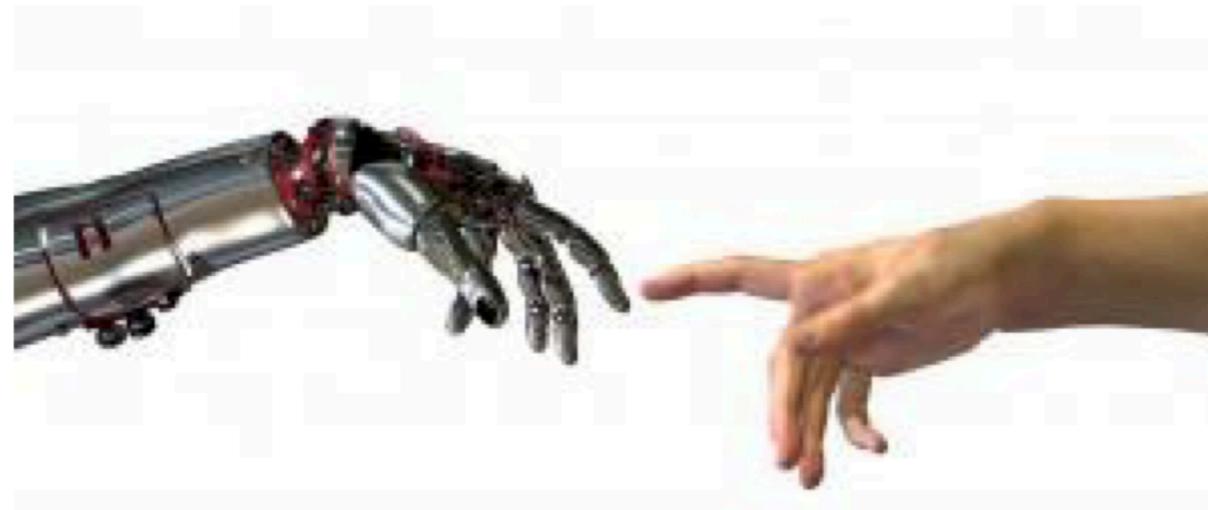


Intellectual Challenge



Figure credit: Hakan Bilen

Intellectual Challenge



- Making machines see
- Extracting semantic information from signals

Table 1

3	120	23	33
6	34	45	56
1	59	67	90
90	99	23	84
200	121	89	55

Algorithms

- Segmentation (Graph partitioning, Non-parametric density estimation)
- Denoising (Sparsity)
- Template Matching
- Deep Neural networks

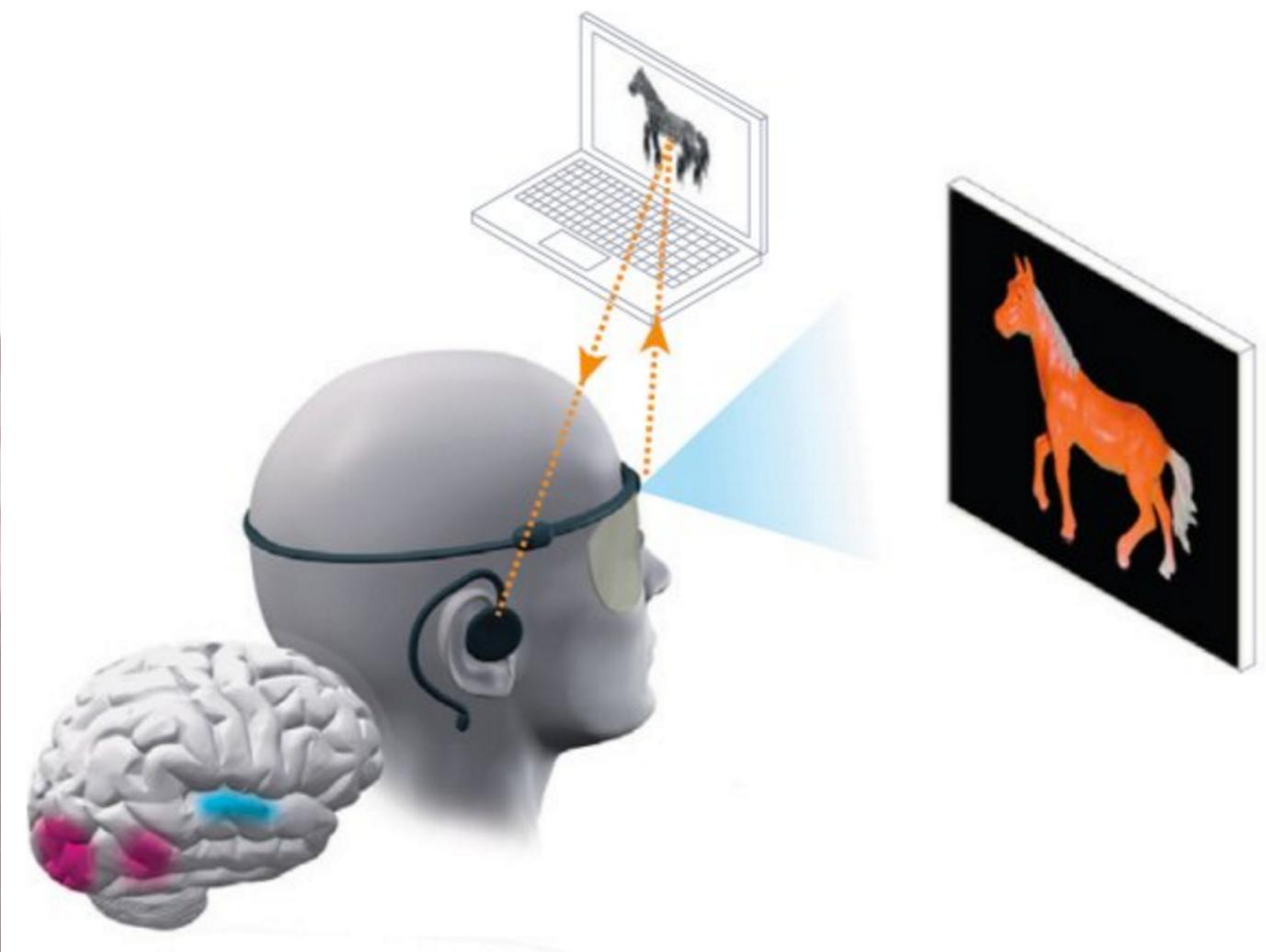
Applications



Self driving cars



Assistive Vision

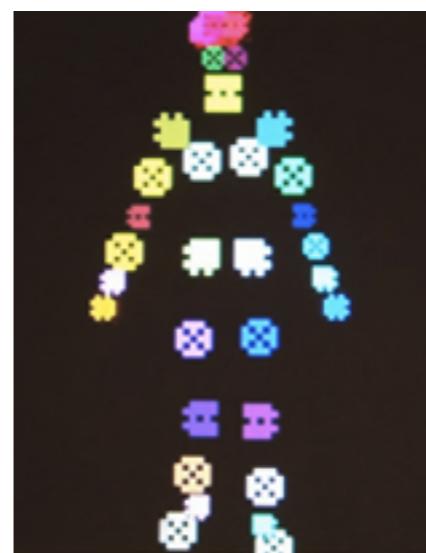
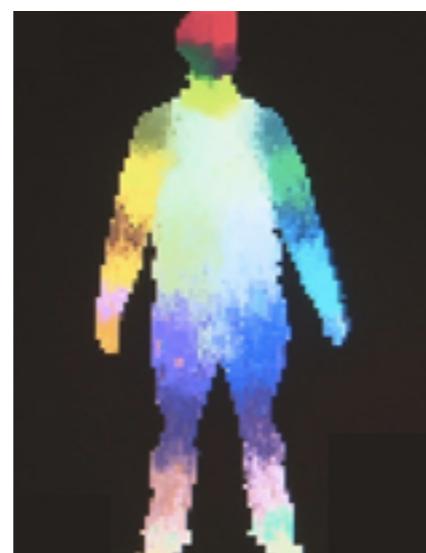


Applications



Surveillance

Applications



Human-Computer Interfaces

Challenges

Challenges



Challenges

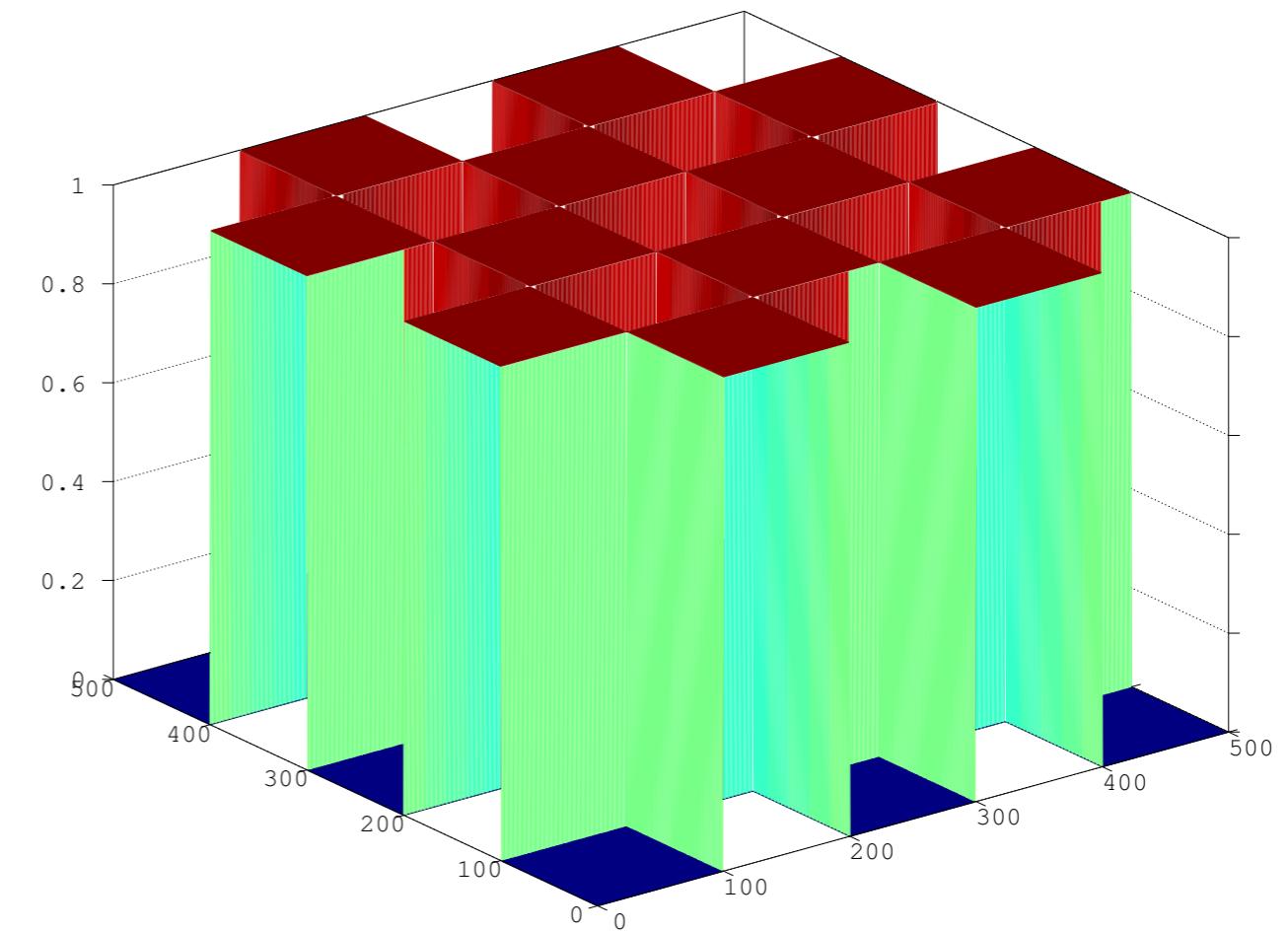
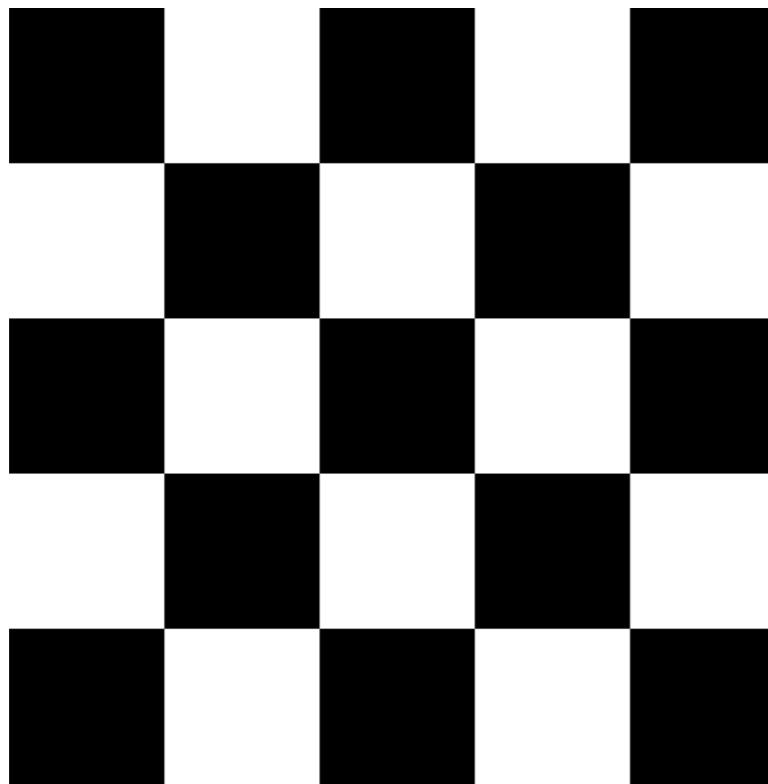


Challenges



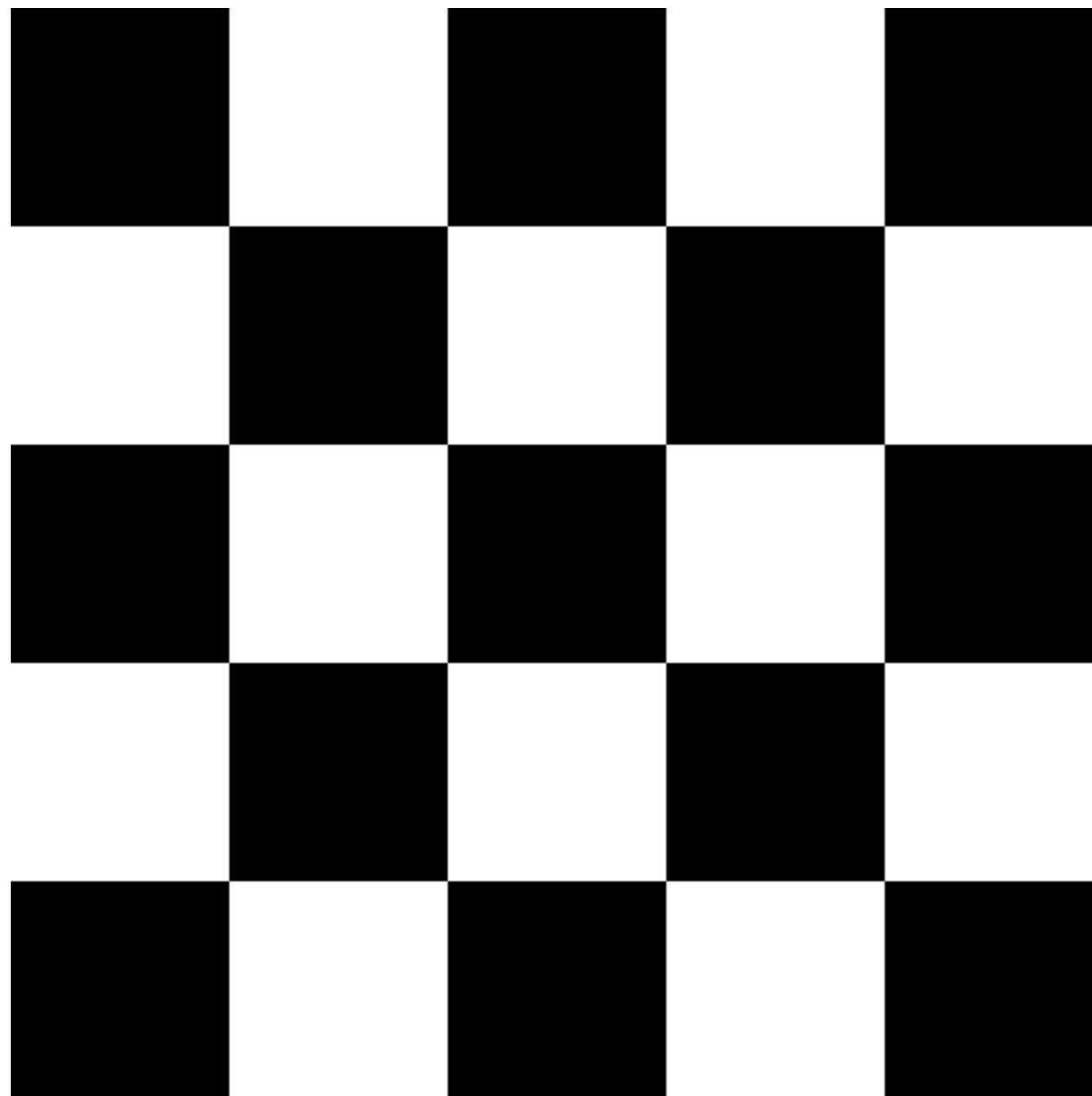
An image

- A rectangular grid of picture elements (pixels).
Each pixel can be a scalar or a vector(3 or 4 element)



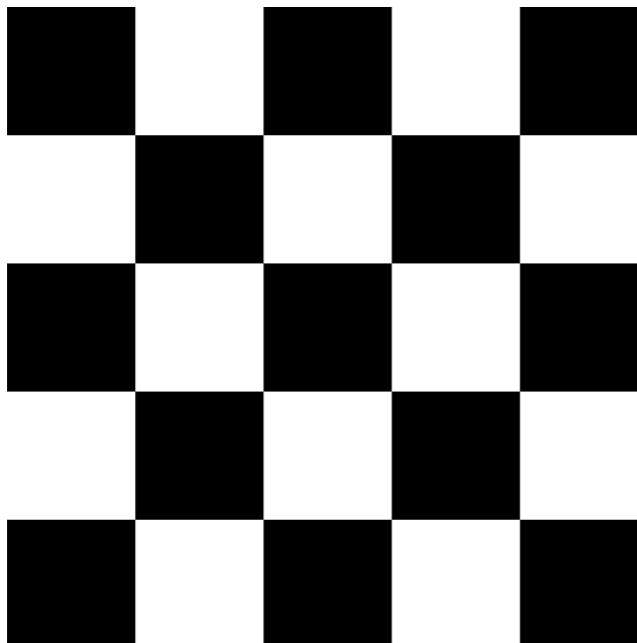


Interpreting an image



Not a very tough task to interpret this image

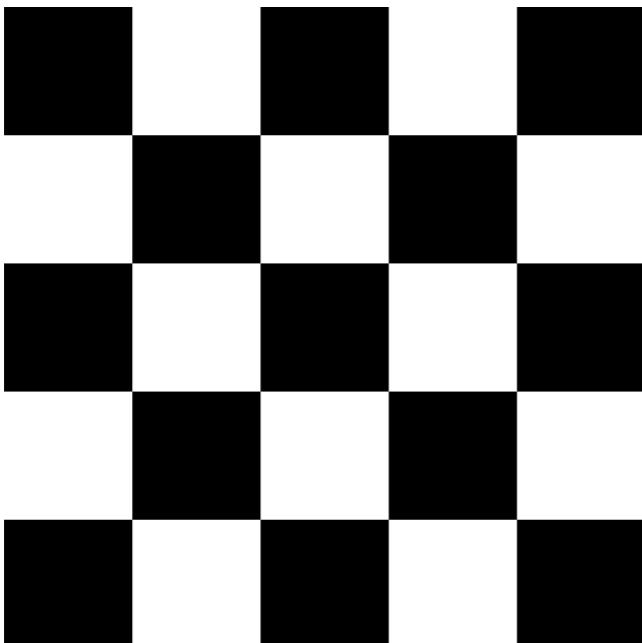
Interpreting an image



However, same checkerboard could be oriented differently, and say extracting the checkerboard pattern would not be trivial with geometric variation

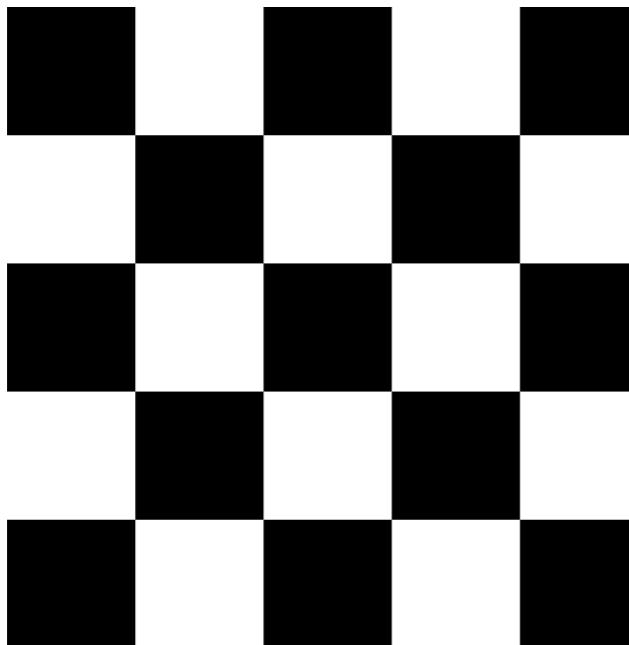


Interpolation



Checkerboard on cookies!

Interpreting an image



Checkerboard
pattern variants

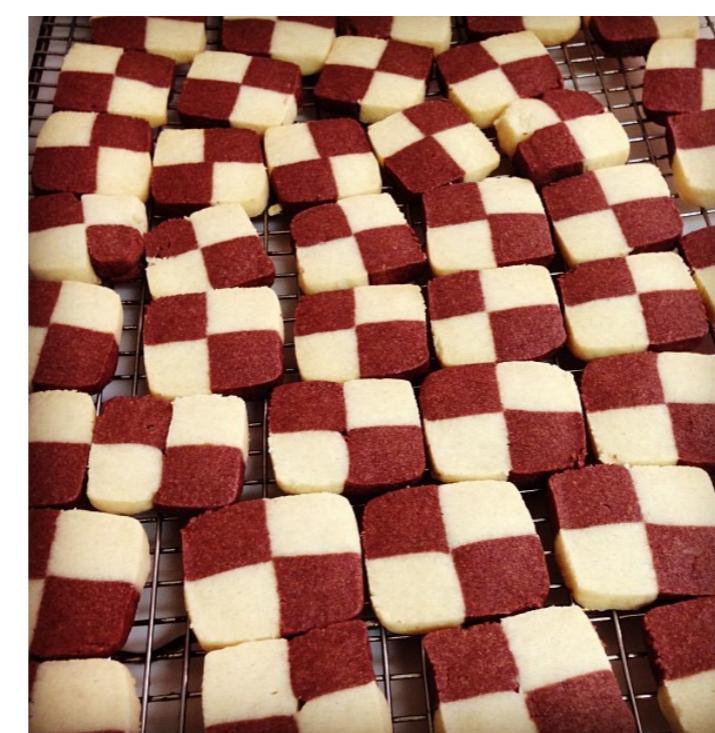
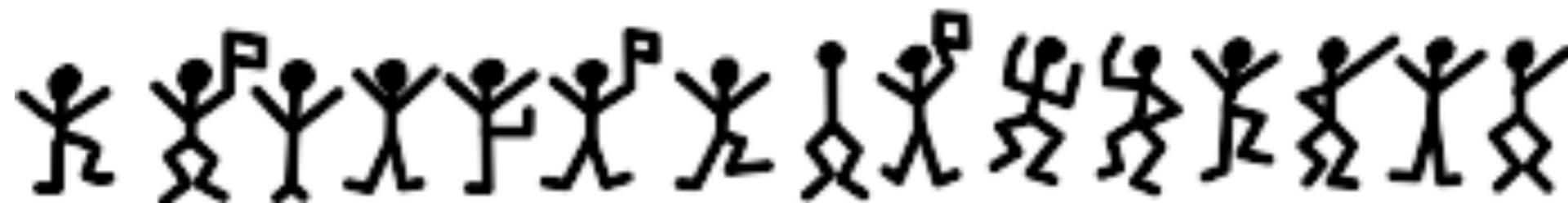


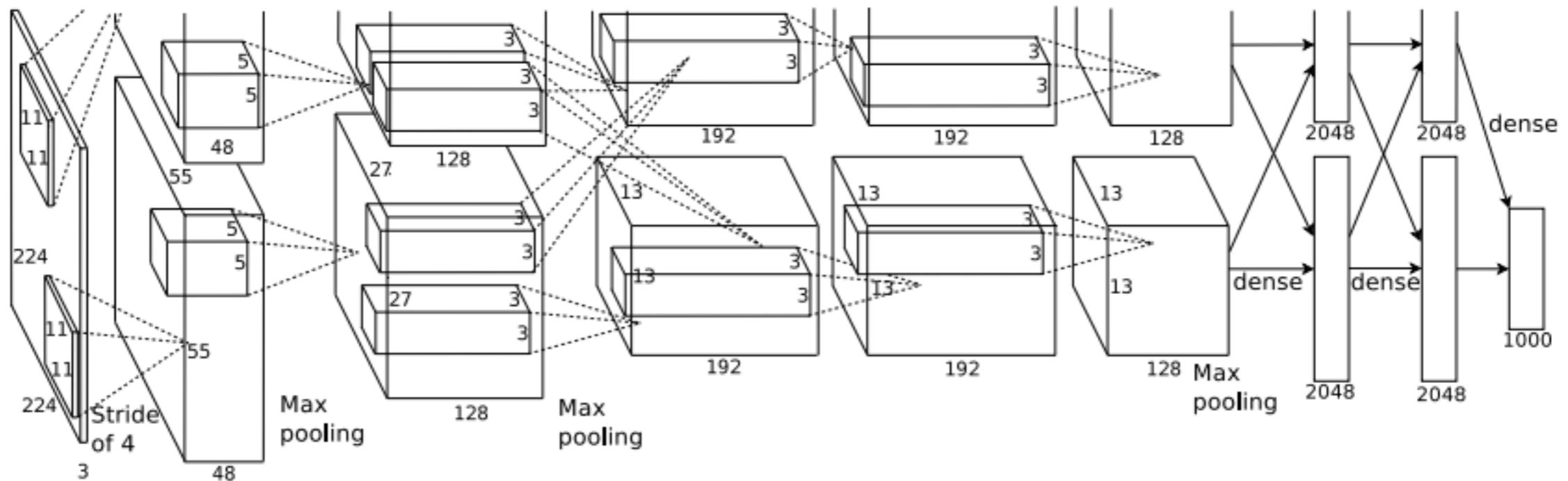


Image interpretation



The Adventure of the dancing men
- Sherlock Holmes by Arthur Conan Doyle

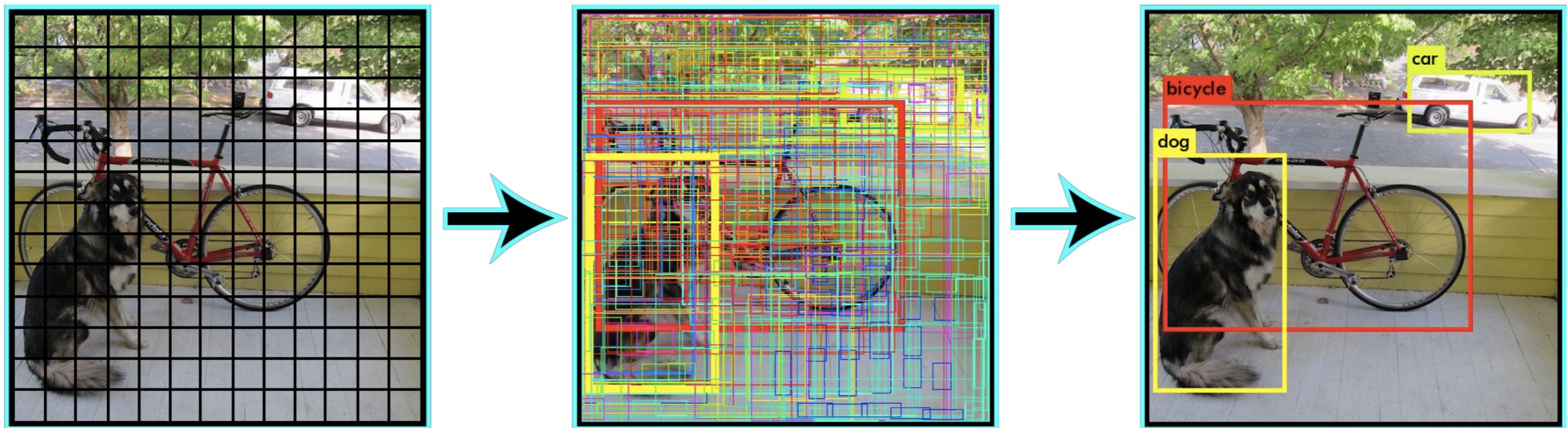
Some Successes



ImageNet Classification with Deep Convolutional
Neural Networks

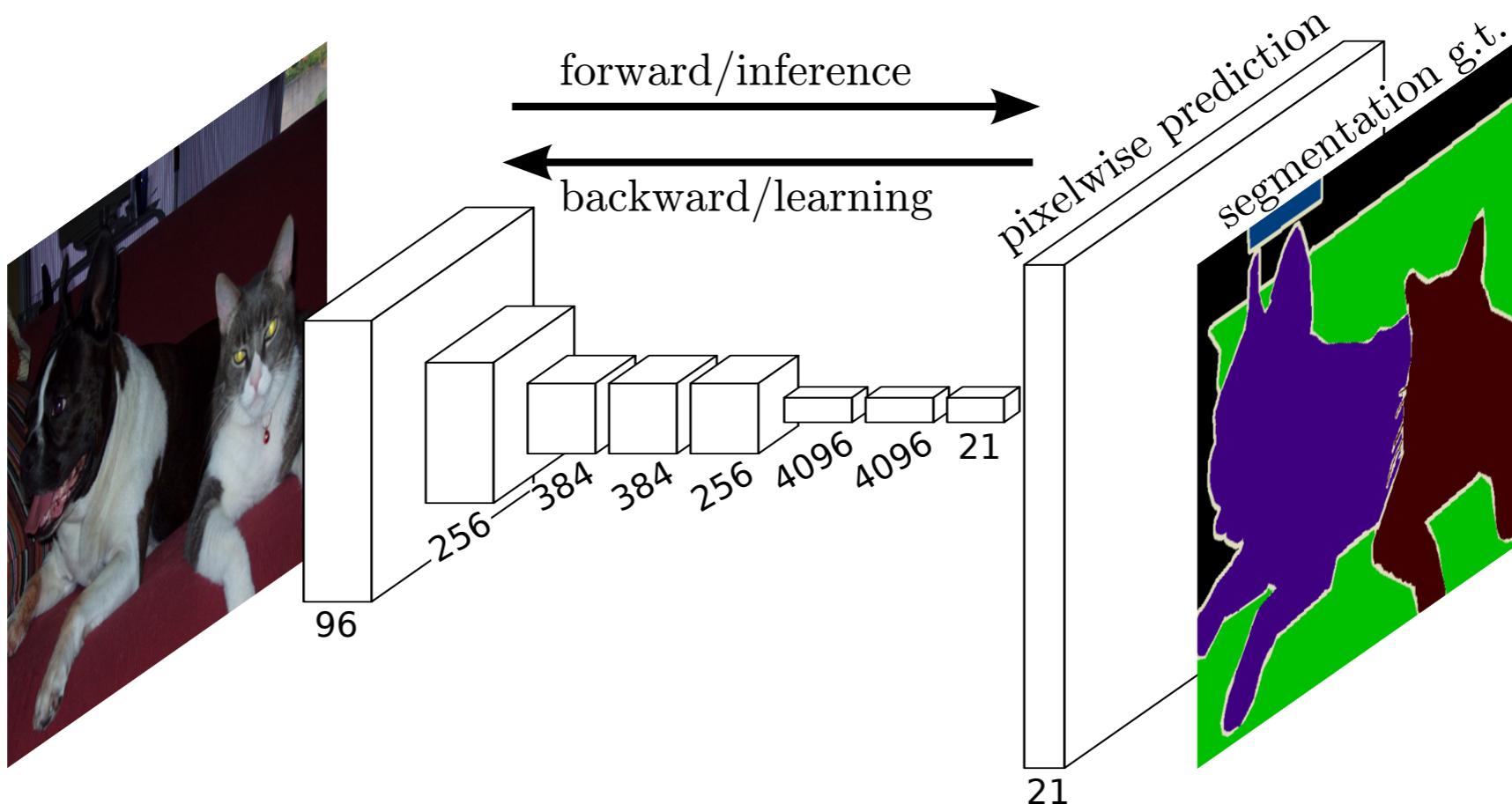
Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton
NIPS 2012

Some Successes



You Only Look Once: Unified, Real-Time Object Detection
Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi
CVPR 2016

Recent Successes

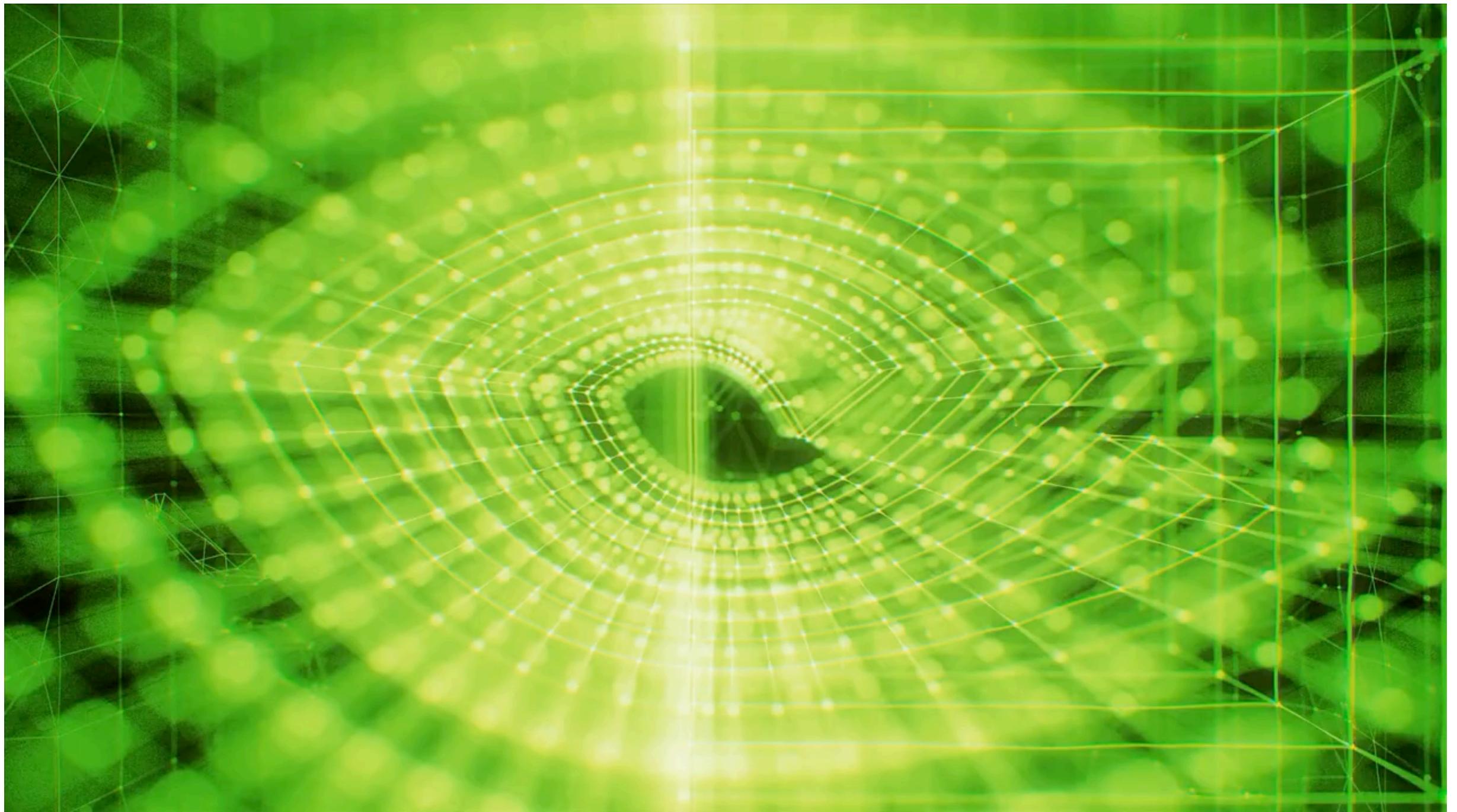


Fully Convolutional Networks for Semantic Segmentation

Jon Long*, Evan Shelhamer*, Trevor Darrell

CVPR 2015

Recent Successes



<https://www.youtube.com/watch?v=kSLJriaOumA>

Recent Successes

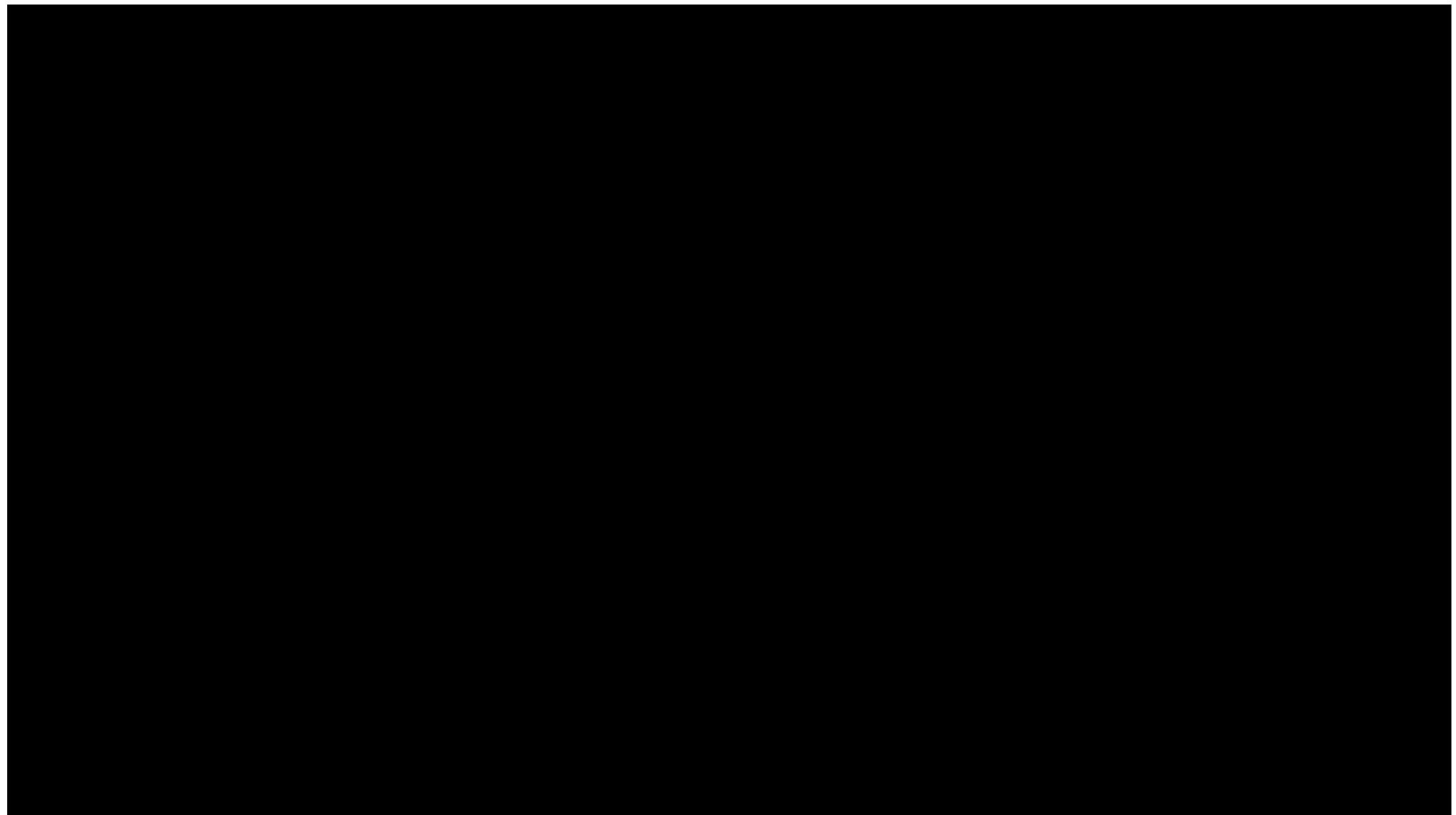


Soccer On Your Tabletop

Konstantinos Rematas
Ira Kemelmacher-Shlizerman
Brian Curless
Steve Seitz

<https://www.youtube.com/watch?v=eRGAB4QBS6U>

Recent Successes



<https://www.youtube.com/watch?v=PCBTZh41Ris>

Conclusion

- Study of visual recognition is one of the classical and interesting problems that is fascinating
- Solving this enables many applications
- This could enable us to move towards real developments in AI