

知能情報論 イントロダクション

2016年4月13日

東京大学 大学院情報理工学系研究科

原田達也

講義概要

- 知能の根元には情報を介した主体と環境とのインタラクションがあり，人間を含めた実環境からの情報をいかに要約し利用するかが知能システム構築のカギとなる．ここでは，その基礎理論，知能システムの原理や仕組み，さらに具体的な応用について論じる．
- 画像認識を例に紹介する．
- An interaction between agent and environment via information is a base of intelligence. Therefore how to summarize and utilize information from the real environment is a key element in building intelligent systems. This class will introduce you to a basis theory of intelligent information processing, and give you experience creating an object recognition system from images.

講義の情報

- WEB
 - <http://www.mi.t.u-tokyo.ac.jp/harada/lectures/IIT/>
 - I D : sum2016 (summer 2016の意味)
 - パスワード : h28-iit
- 注意 : シラバスのスケジュールは前後することがある. また急遽予定を変更する可能性もあるので, 掲示板と上記WEBを良く注意しておくこと.

成績

- 出席 (30-50%)
- レポート (50-70%) : プログラミング課題

Schedule

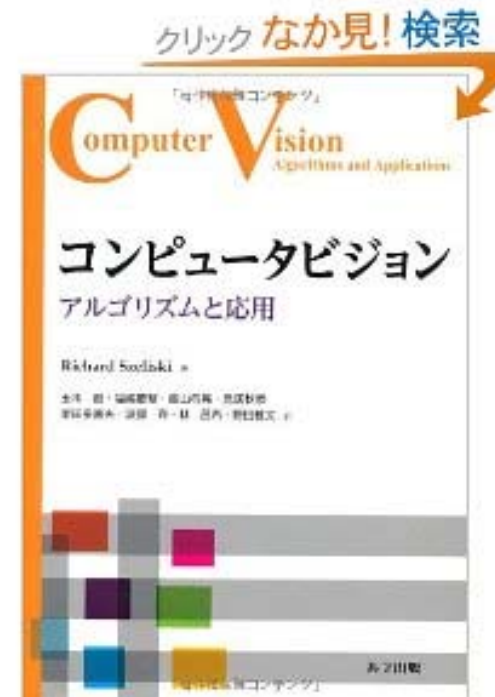
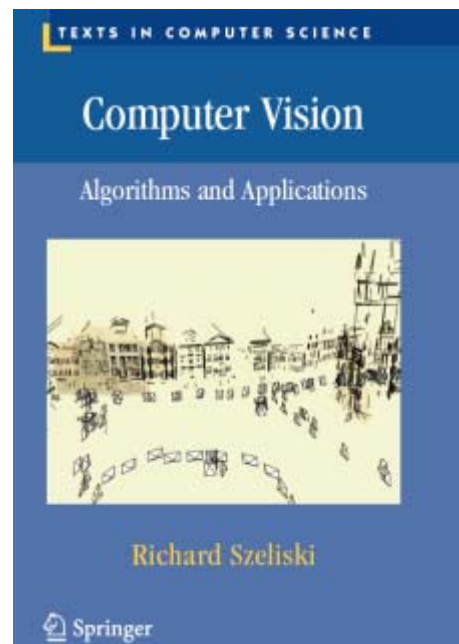
- 4/13: Introduction
- 4/20: Short history, What's image recognition?
- 4/27: Detectors and descriptors
- 5/04: Golden week holidays!
- 5/11: Statistical feature extraction
- 5/18: Image coding and pooling
- 5/25: Classifiers
- 6/01: 講義は実施しない
- 6/08: Deep learning and convolutional neural networks
- 6/15: Object detection
- 6/22: Caption generation (Prof. Ushiku)
- 6/29: Cancel (CVPR)
- 7/06: Image and language (Prof. Ushiku)
- 7/13: Instance recognition and search

必要な知識

- 必要な知識は適宜説明する
 - 線形代数
 - 多変量解析
 - 確率・統計
 - パターン認識
 - 機械学習
 - 情報理論

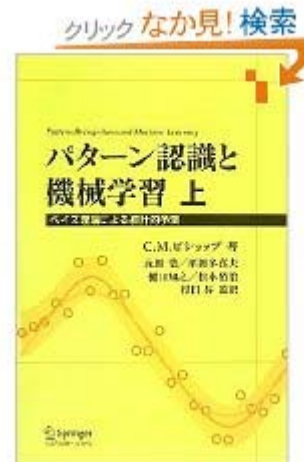
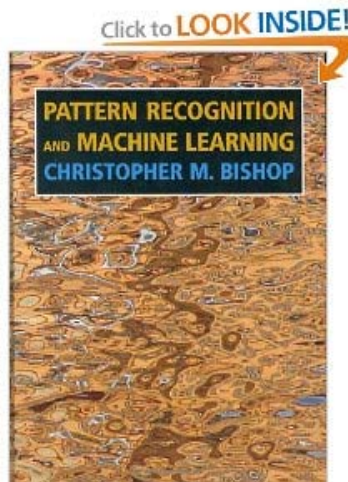
参考書

- Richard Szeliski. Computer Vision: Algorithms and Applications (Texts in Computer Science). Springer-Verlag.
- Richard Szeliski (著), 玉木 徹 (翻訳), 福嶋 慶繁 (翻訳), 飯山 将晃 (翻訳), 鳥居 秋彦 (翻訳), 栗田 多喜夫 (翻訳), 波部 斉 (翻訳), 林 昌希 (翻訳), 野田 雅文 (翻訳). コンピュータビジョン—アルゴリズムと応用—. 共立出版.

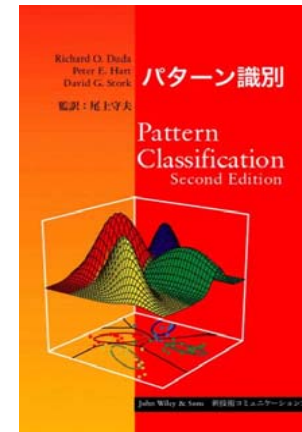
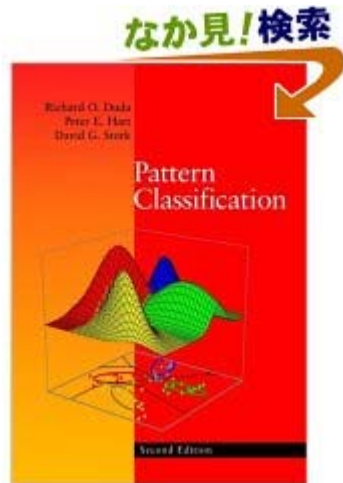


参考書

- Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag.
- C. M. ビショップ 著, 元田浩, 栗田多喜夫, 樋口知之, 松本裕治, 村田昇 監訳. パターン認識と機械学習 上・下 - ベイズ理論による統計的予測. シュプリンガー・ジャパン株式会社.



参考書



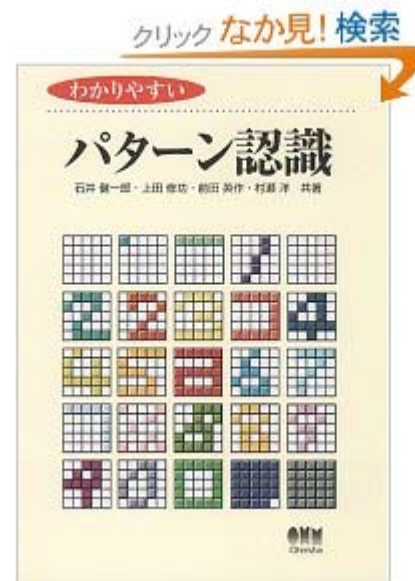
- Pattern Classification

- Richard O. Duda (著), Peter E. Hart (著), David G. Stork (著)
- ハードカバー: 654ページ
- 出版社: Wiley-Interscience; 2 Sub版 (2000/10)
- 言語 英語, 英語, 英語
- ISBN-10: 0471056693
- ISBN-13: 978-0471056690
- 発売日: 2000/10

監修: 尾上守夫 (東大名誉教授)
翻訳: 江尻公一 ほか
編集: アドコム・メディア(株)
定価: 10,500円 (本体10,000)

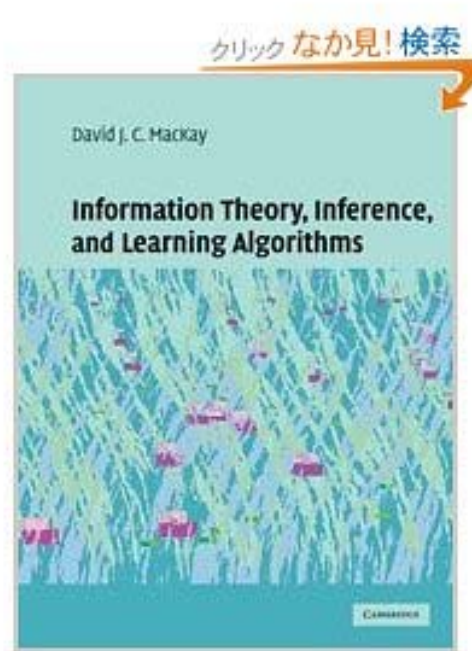
参考書

- 石井 健一郎, 前田 英作, 上田 修功, 村瀬 洋.
わかりやすいパターン認識. オーム社, 1998.



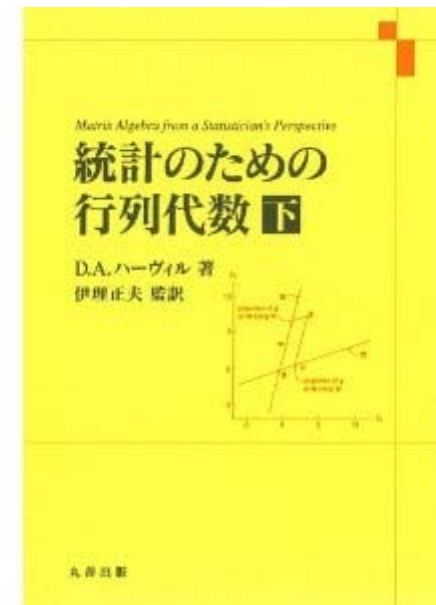
参考書

- David J. C. MacKay. Information Theory, Inference, and Learning Algorithms. Cambridge University Press.



参考書

- D. A. ハーヴィル. 統計のための行列代数 上下
・ 丸善出版.



- The Matrix Cookbook
– <http://orion.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf>

教科書

- 講談社，機械学習プロフェッショナルシリーズ「画像認識」，原田達也
- 本年度，執筆予定

これは何でしょう？



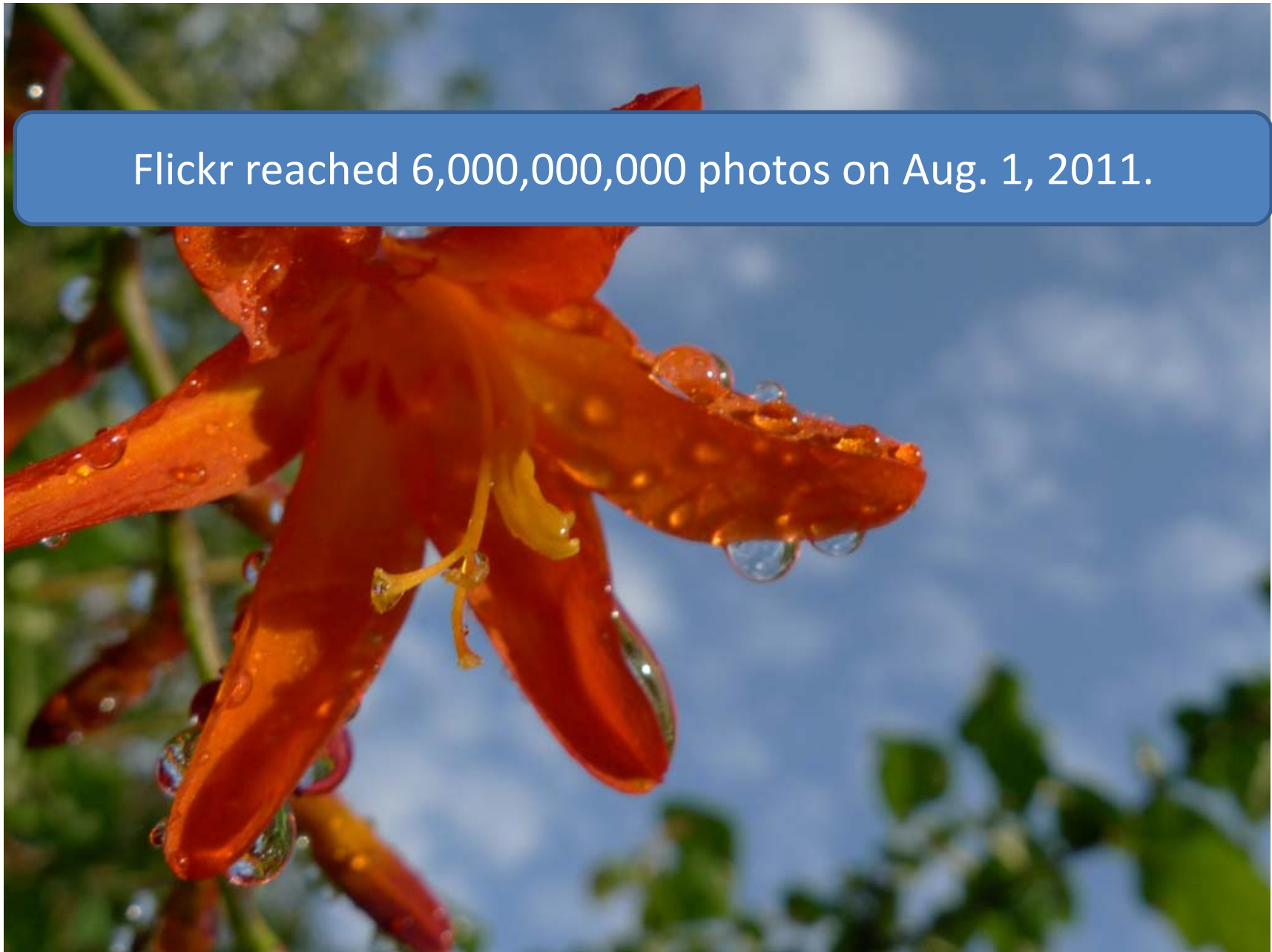
beach, water, people, kauai, tree



ocean, coral, fish, angelfish, reefs

実世界認識知能の構築

Flickr reached 6,000,000,000 photos on Aug. 1, 2011.

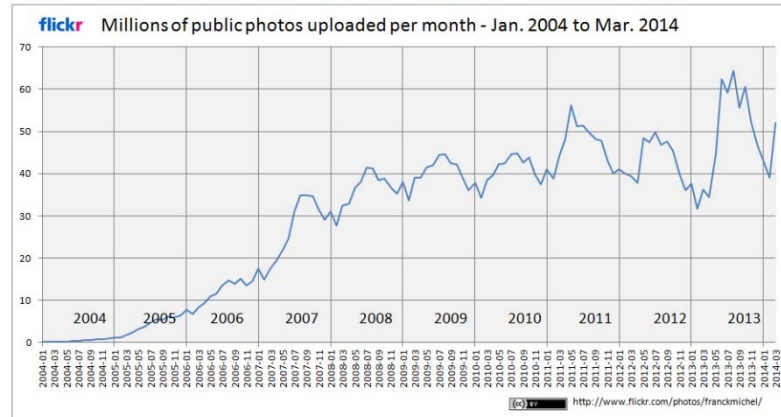


<http://www.flickr.com/photos/eon60/6000000000/>

Growth of Multimedia Data

Image

How many photos are uploaded to Flickr every day, month, year?



Video

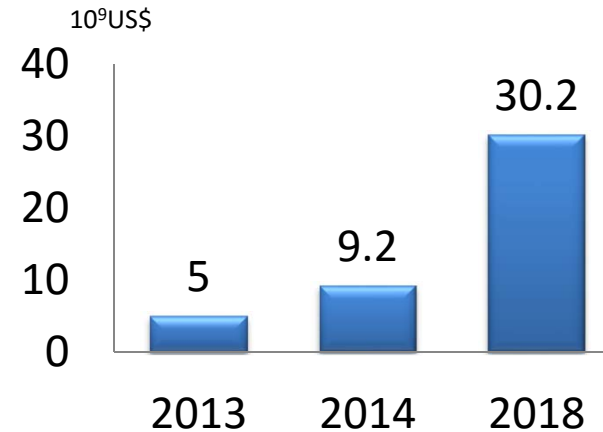


100 hours of video are uploaded to YouTube every minute

<https://www.youtube.com/yt/press/statistics.html>

Lifelog

Wearable Computing Market



Samsung Gear Live



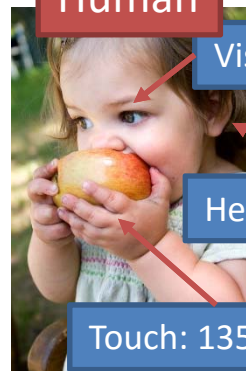
Google Glass



Nest

BCC Research, Wearable Computing: Technologies, Applications and Global Markets, 2014.

Human

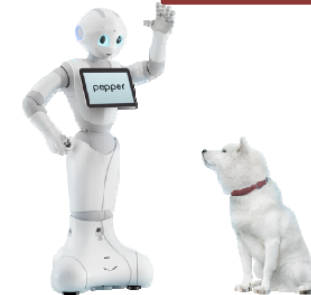


Vision: 8.96 Mb/s

Hearing: 30 Mb/s

Touch: 135 mb/s

Robot

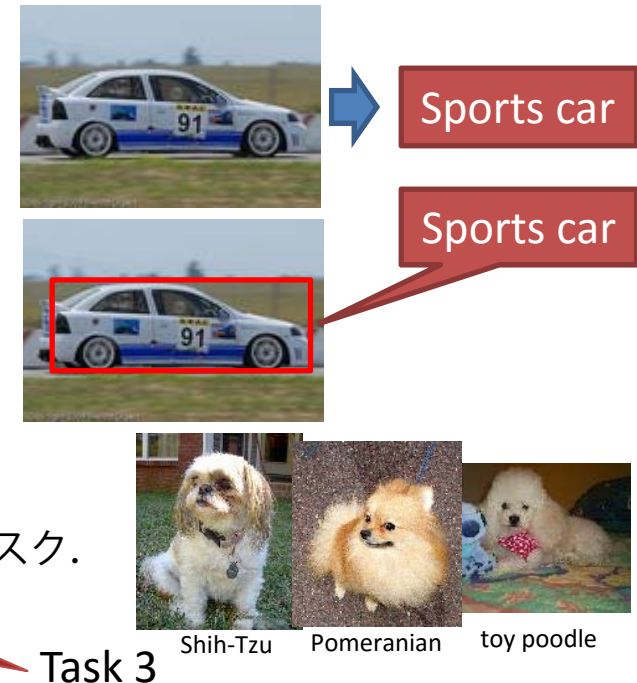


<http://www.softbank.jp/robot/>

A massive amount of multimedia data is uploaded every moment!

大規模データを利用した画像認識

- ILSVRC (ImageNet Large Scale Visual Recognition Challenge)
 - 大規模なデータを利用した, 国際的画像認識のコンペティション
 - <http://www.image-net.org/challenges/LSVRC/2012/index>
 - 現在最も困難な画像認識タスク
- Task 1
 - 120万枚の画像を学習して, 1000クラスの画像を識別
- Task 2
 - 画像内に1000クラスの物体がどこあるのか検出
- Task 3
 - 120の犬の種類を当てるTask 1より分類が困難な識別タスク.



Team	Flat Error
1) SuperVision Univ. of Toronto	0.153
2) ISI (ours) Univ. of Tokyo	0.262
3) OXFORD_VGG Univ. of Oxford	0.270

Team	mAP
1) ISI (ours) Univ. of Tokyo	0.323
2) XRCE/INRIA Xerox Research Centre Europe/INRIA	0.310
3) Uni Jena Univ. Jena	0.246

最新の画像識別性能

multi-model results

	Team	Top 5 error [%]	
post-competition	Google	4.9	Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift
	MSRA, PReLU-nets	4.94	Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification
	Human	5.1	
	Baidu	5.98	
	VGG (arXiv v5)	6.8	
in competition ILSVRC 14	GoogLeNet	6.66	
	VGG (Oxford)	7.32	
	MSRA, SPP-nets	8.06	

人の能力を
超えた？


Baidu (百度) による不正

SECTIONS

HOME

SEARCH

The New York Times




Virtual Reality Headsets Raise Very Real Concerns

TECHNOLOGY


BITS BLOG

Elon Musk Wants Your Hyperloop Ideas



BITS BLOG


IBM Invests to Help Open-Source Big Data Software — and Itself




Jack Dorsey: Twitter as (and Quips

Pure Innovation

LEARN HOW »





Search Bits

ARTIFICIAL INTELLIGENCE

Baidu Fires Researcher Tied to Contest Disqualification

By JOHN MARKOFF JUNE 11, 2015 4:17 PM [Comment](#)

Email

Share

Tweet

Save

More

A week after researchers at Baidu were barred from competing in an international computer vision contest after contest organizers discovered that scientists at the company broke the rules for the competition, the Chinese web service said that it had fired the team leader.

The company did not name the individual in a [brief posting Thursday](#) morning on the company's website, but a spokesman said that Ren Wu, a distinguished scientist at the company's Institute of Deep Learning, was no longer working for Baidu.

PREVIOUS POST

Uber Releases Video Game Drivers

Visit the **Techn** the industry. »

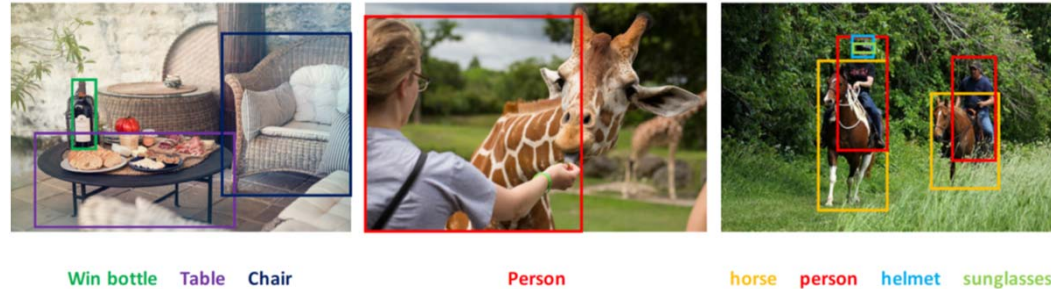
MOST VIEWED

1. Google Inter Reality Devi
2. In Busy Silic Demand
3. Sheryl Sand Meditations

http://bits.blogs.nytimes.com/2015/06/11/baidu-fires-researcher-tied-to-contest-disqualification/?_r=1

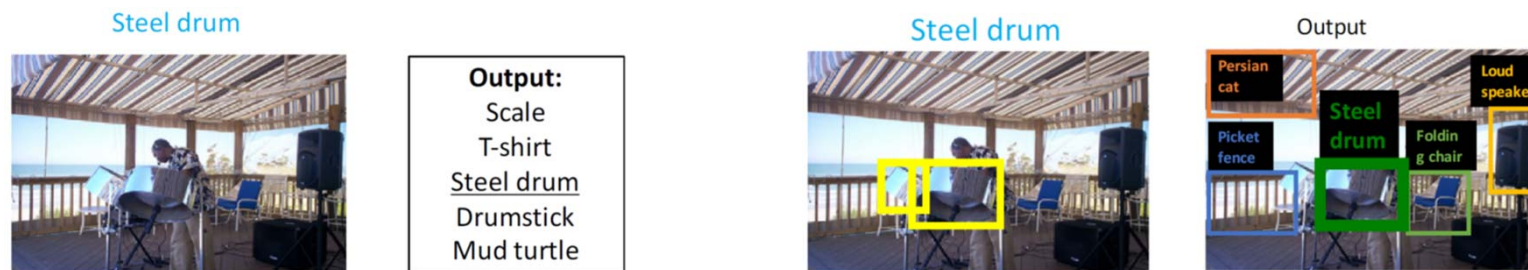
ILSVRC2015

- Main competitions
 - Object detection
 - 200 fully labeled categories and 0.4 million images



http://image-net.org/challenges/talks/ILSVRC+MSCOCO_12_17_15_detection.pdf

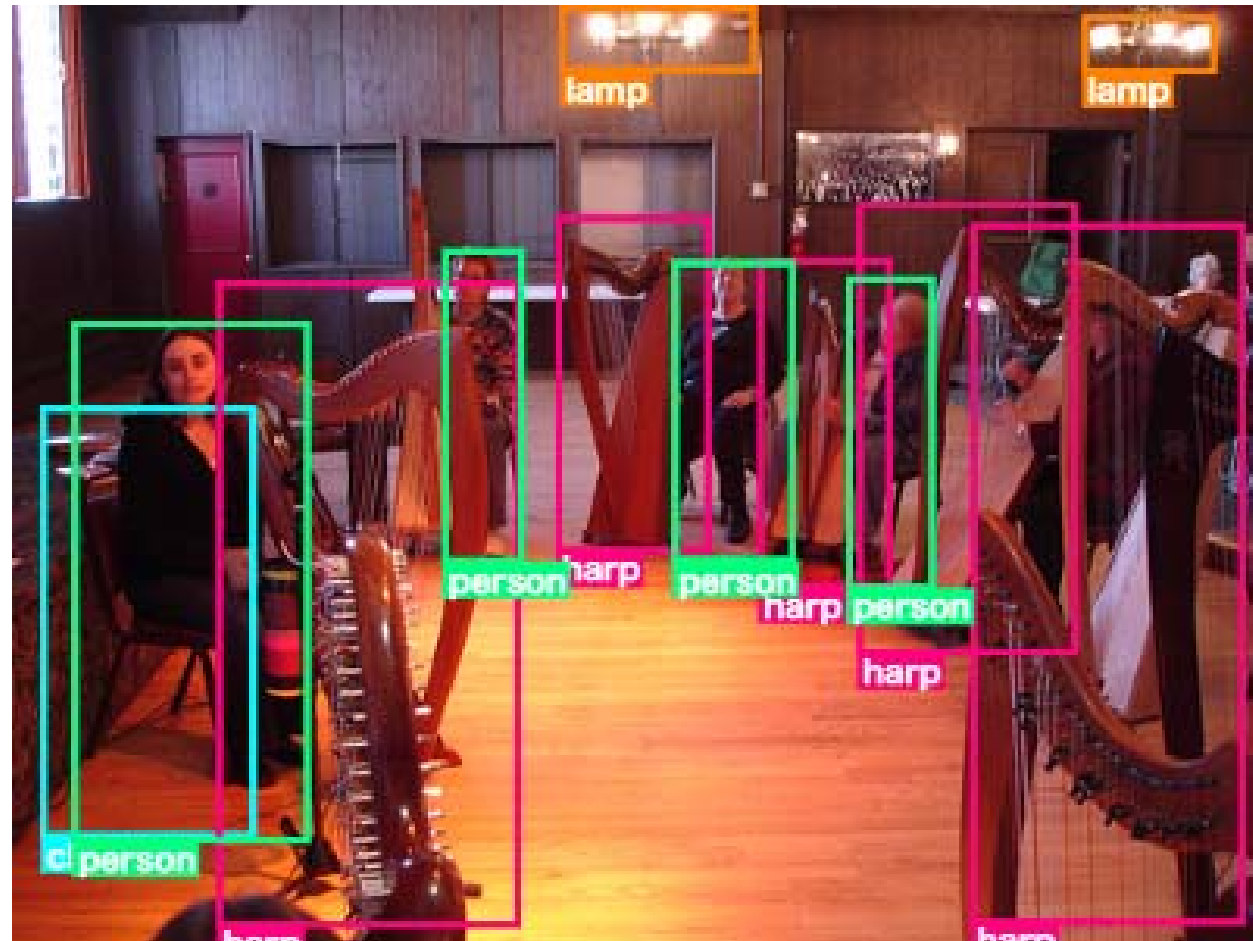
- Object classification and localization
 - 1000 categories and 1.2 million images



http://image-net.org/challenges/talks/ILSVRC2015_12_17_15_clsloc.pdf

- Taster competitions
 - Object detection from video
 - Fully annotated 30 object classes across 5,354 snippets
 - Scene classification
 - 401 scene categories, 8.1M train, 20k val, 381k test

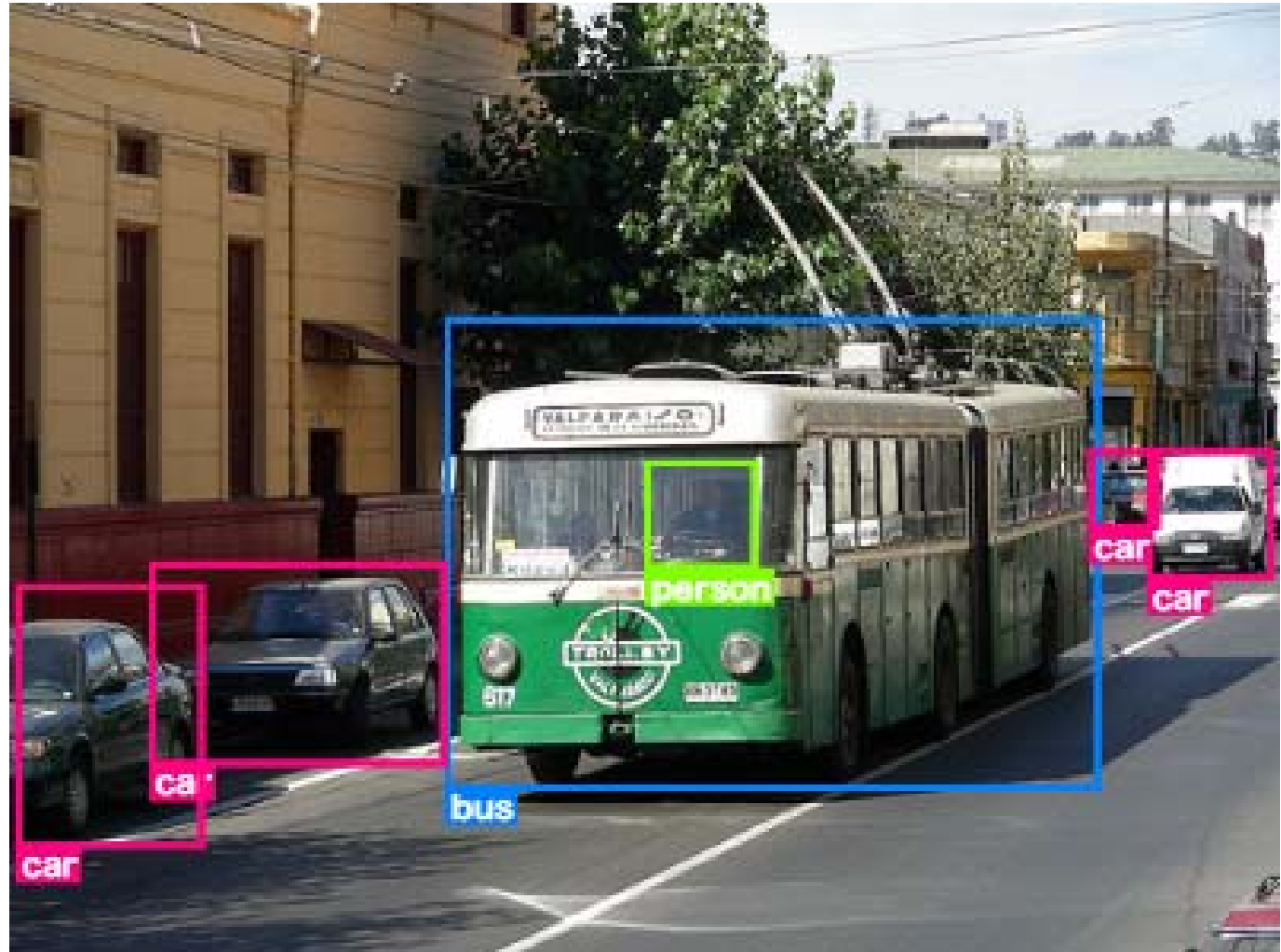
Good
result



(harp, person, chair, lamp)

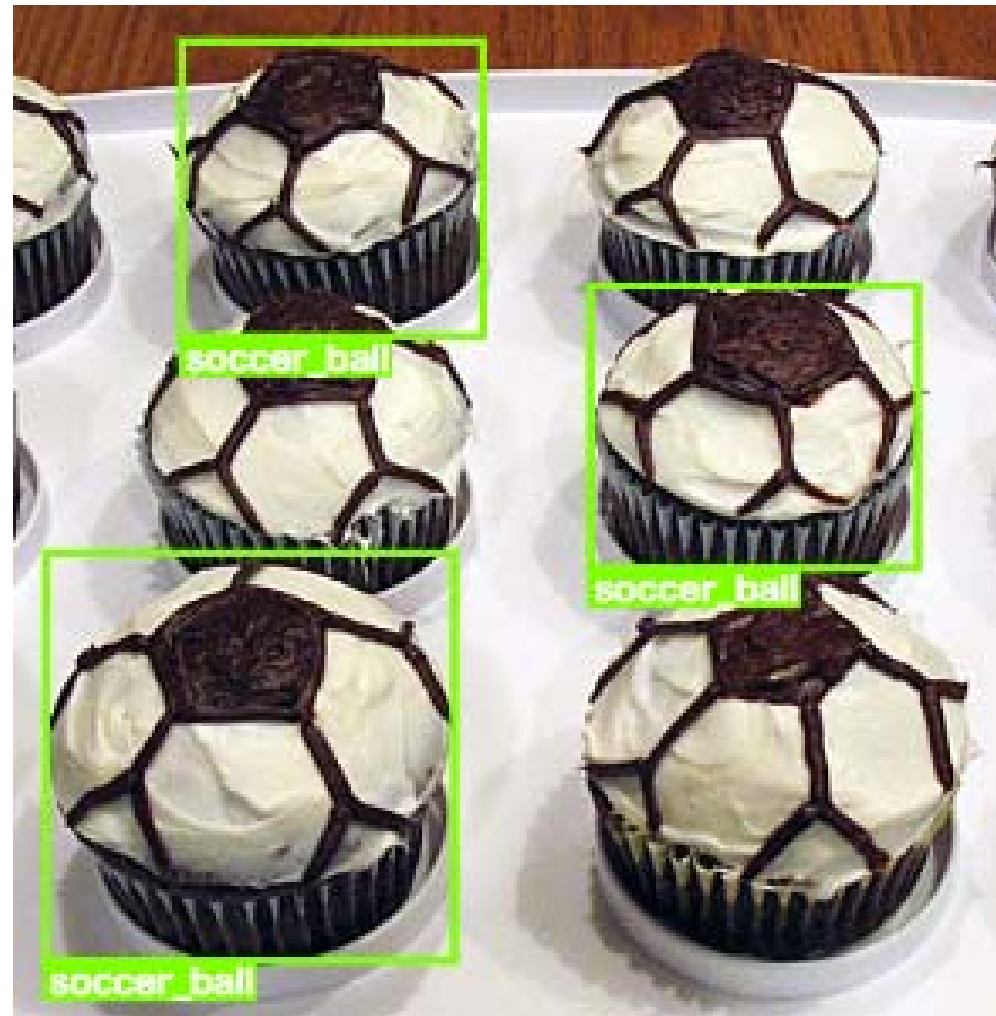
Good
result

(car, bus, people)

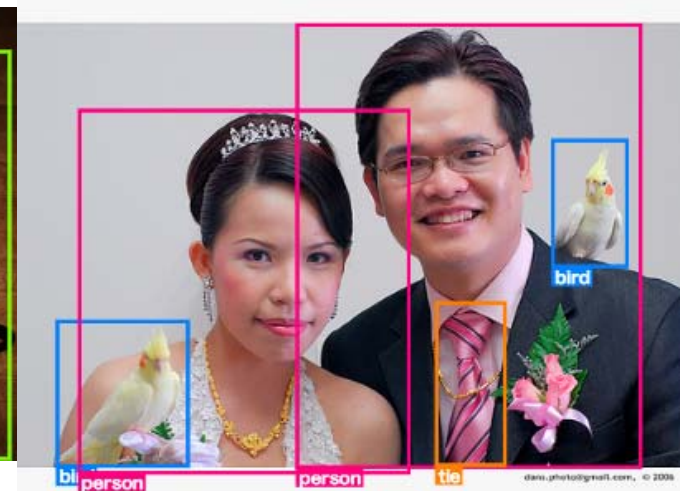
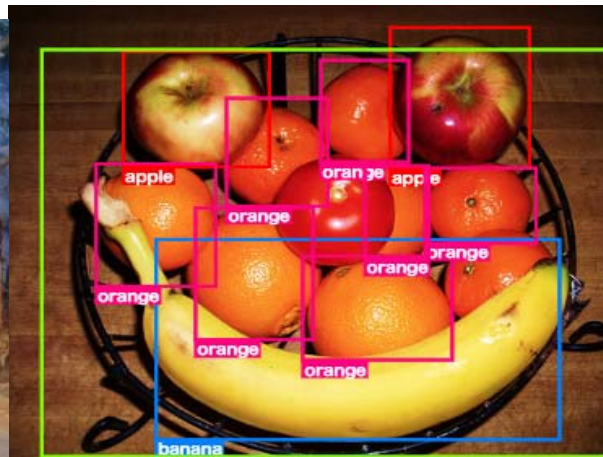
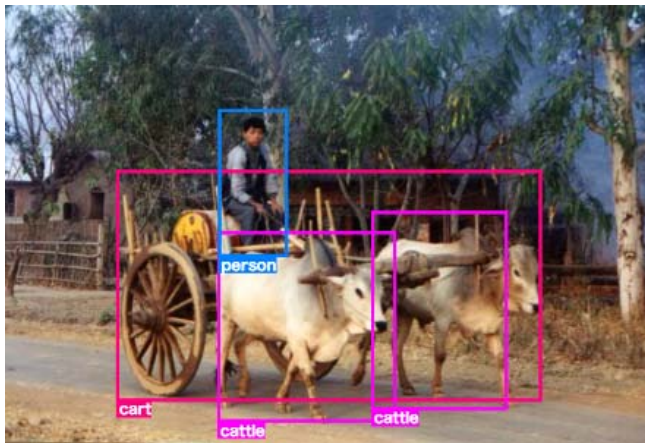
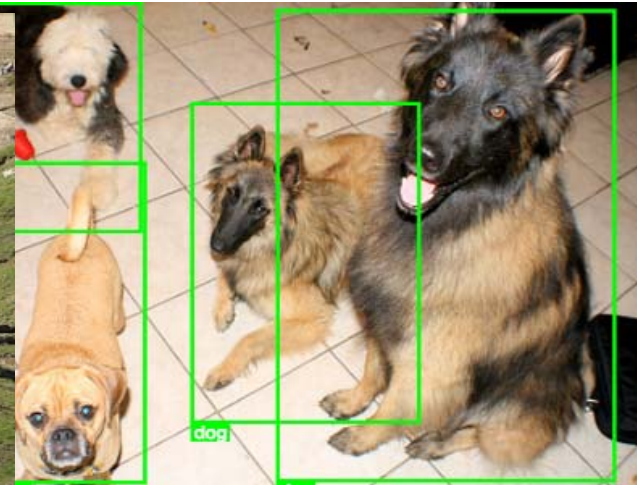
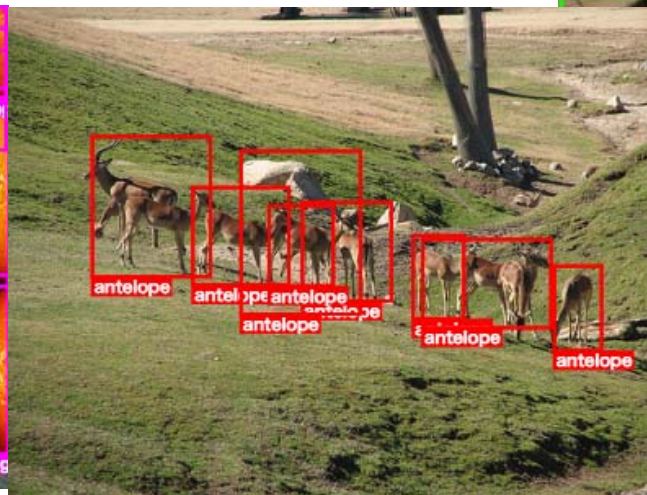
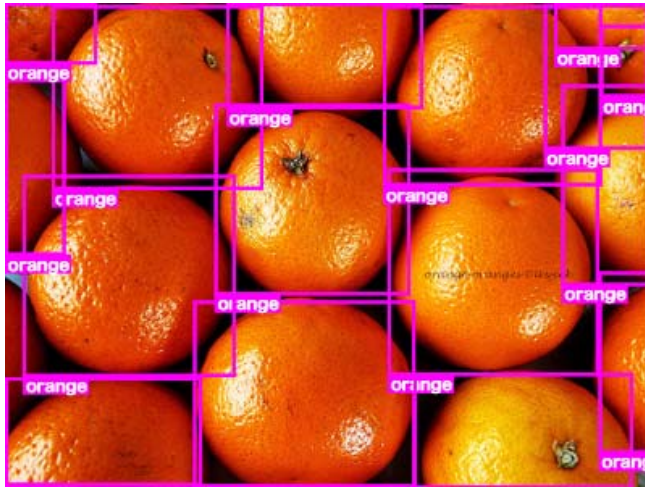


Bad
result

(soccer ball)



Another result



実世界応用 1

人工知能ゴーグルの開発

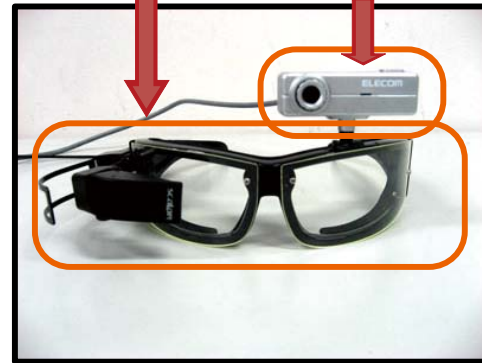
- 提案手法の実世界応用：人工知能ゴーグル
 - 身の回りの物体の素早い認識・検索を実現
 - HMDによる情報提示，記憶支援（忘れ物検索）

Head Mount Display

displays labels of objects and scene with the image.

Camera

captures images, which the wearer is seeing now.



Portable Computer

recognizes the images quickly that the camera is capturing, and shows the results on HMD. Moreover, it accumulates the images and labels, and enables the wearer to search those images by labels.



Project Glass



<https://plus.google.com/111626127367496192147/posts>

AI Goggles on iPhone Project

Head Mount Display

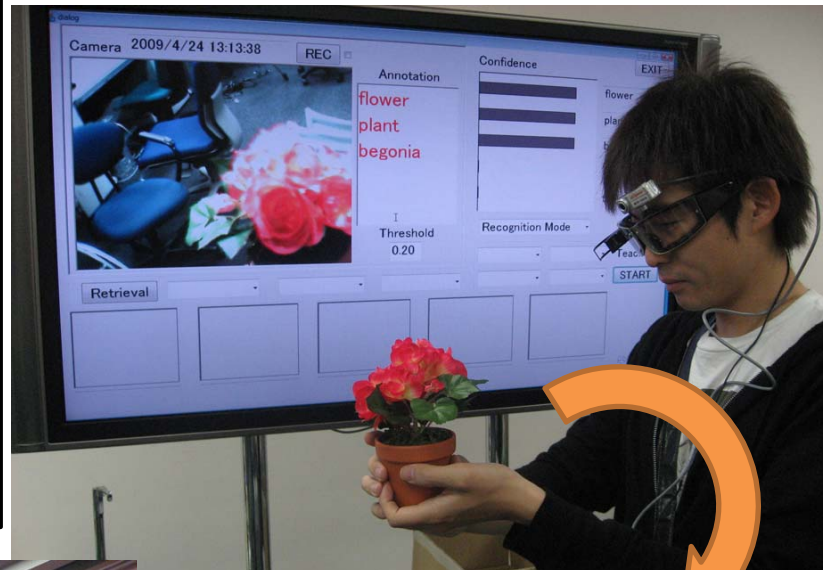
displays labels of objects and scene with the image.

Camera
captures images, which the wearer is seeing now.

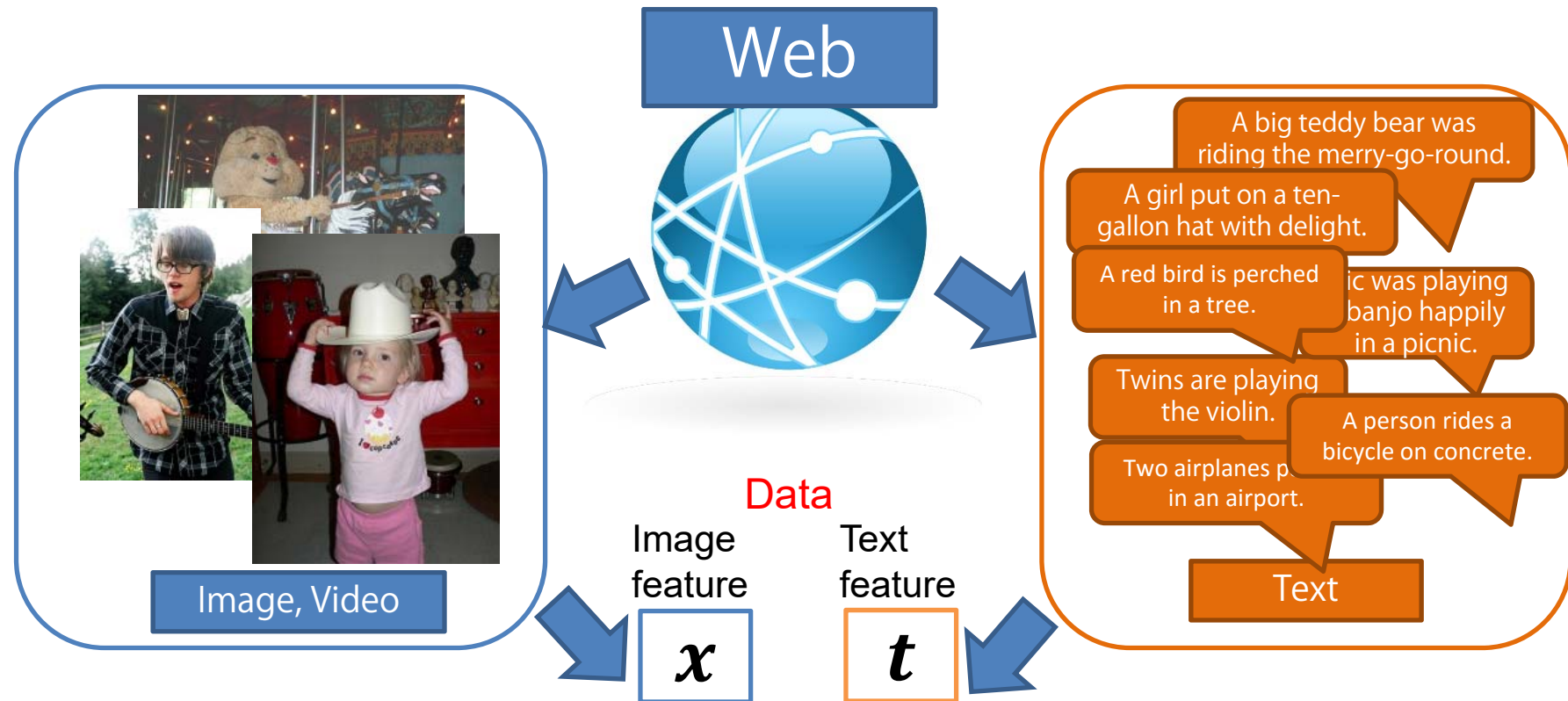


Portable Computer

recognizes the images quickly that the camera is capturing, and shows the results on HMD. Moreover, it accumulates the images and labels, and enables the wearer to search those images by labels.



Overview: Machine Learning for Visual Recognition



Learning the relationships between images and text

Mapping function

Risk

$$R(\mathbf{w}) = \iint \underbrace{r(\Psi(\mathbf{x}, \mathbf{w})|\mathbf{t})}_{\text{Loss}} \underbrace{p(\mathbf{x}, \mathbf{t})}_{\text{Joint probability}} d\mathbf{x} d\mathbf{t}$$

Loss

Joint probability

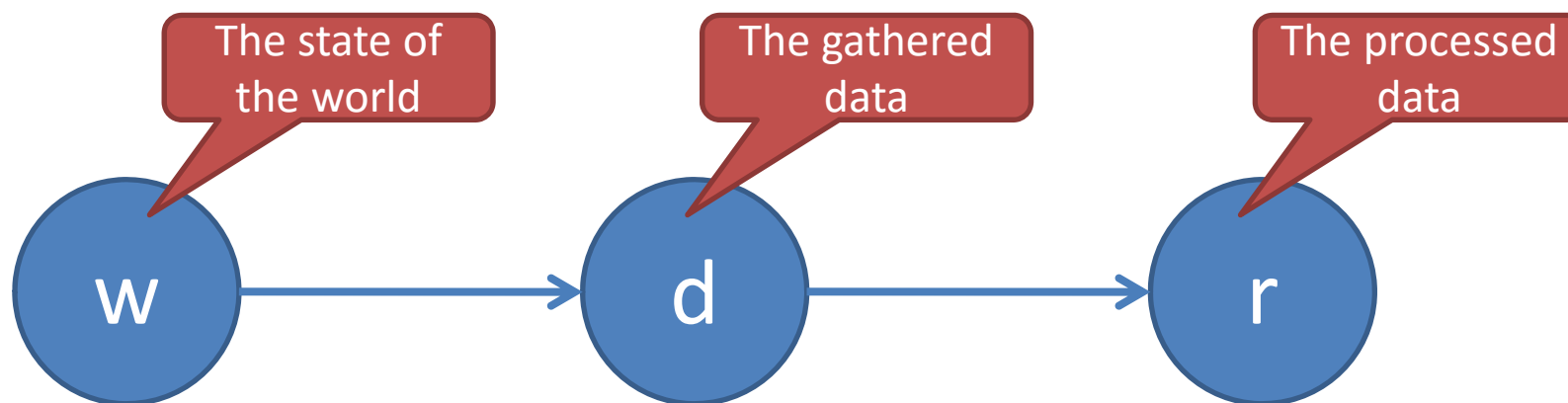
Minimization



Ex) SGD

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \epsilon \nabla_{\mathbf{w}} r(\Psi(\mathbf{x}, \mathbf{w}_t)|\mathbf{t})$$

The data processing theorem



Markov chain

$$P(w, d, r) = P(w)P(d | w)P(r | d)$$

The average information

$$I(W; D) \geq I(W; R)$$

The data processing theorem states that data processing can only destroy information.

知識はどこから降ってくる？

- 膨大な情報を利用
 - インターネット上の大規模画像
 - リッチな画像表現，効率な線形識別機の学習
- 人に尋ねる
 - クラウドソーシング
 - 能動学習
- 対象とは別の知識を活用
 - 転移学習，ドメイン適合，マルチタスク学習
 - ゼロショット学習
 - アトリビュート





数学の問題です。辺AB、BC、CAの長さがそれぞれ12、11、10の三角形ABCを考える。∠A...

aicezukiさん

数学の問題です。

辺AB、BC、CAの長さがそれぞれ12、11、10の三角形ABCを考える。∠Aの二等分線と辺BCの交点をDとすると、線分ADの長さを求めよ。

解答だけでなく途中計算もよろしくお願いいたします。

()

質問日時: 2011/2/25 13:37:13

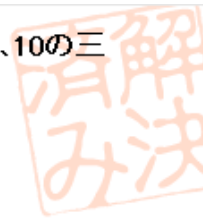
解決日時: 2011

回答数: 1

お礼: 100枚

閲覧数: 99,931

ソーシャルブック



クラウドソーシング



ベストアンサーに選ばれた回答

jawtdpgさん



ADは∠Aの二等分線なので

$$BD:DC=AB:CA=12:10=6:5$$

$$BC=11よりBD=11*6/11=6$$

△ABCにおいて

$$\cos B = (AB^2 + BC^2 - CA^2) / (2 * AB * BC)$$

$$= (144 + 121 - 100) / 264$$

$$= 65 / 264 = 5/8$$

△ABDにおいて

$$AD^2 = AB^2 + BD^2 - 2 * AB * BD * \cos B$$

$$= 144 + 36 - 144 * 5/8$$

$$= 180 - 90 = 90$$

より

$$AD = 3\sqrt{10}$$

だと思いますがいかがでしょうか？

http://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1256265042

回答日時: 2011/2/25 14:09:44

違反報告

Positive and Unlabeled Problem



the obtained dataset has properties by which (1) assigned labels are definitely positive and (2) absent labels are not necessarily negative.



label
Cat
Wall
House



label
Bed
TV
Room



label
Human
Bicycle
Road