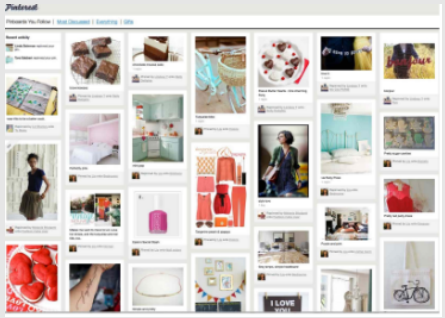
Deep Learning for Medical Images

Nicolas Thome
Université Pierre et Marie Curie (UPMC)
Laboratoire d'Informatique de Paris 6 (LIP6)

Context

Big Data: Images & videos everywhere

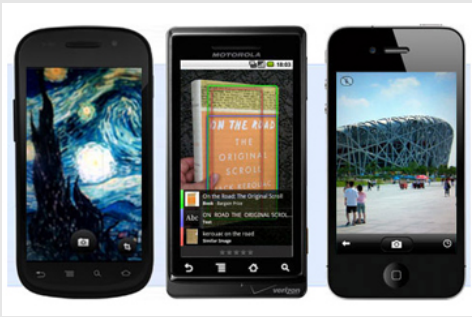
BBC: 2.4M videos

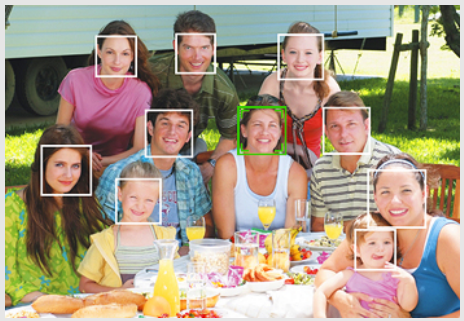


Facebook: 140B images

100M monitoring cameras

- Obvious need to access, organize, search, or classify these data: **Semantic Annotation of Visual Data**
- Huge number of applications: mobile visual search, robotics, autonomous driving, augmented reality, medical imaging *etc*
- Leading track in major CV conferences during the last decade





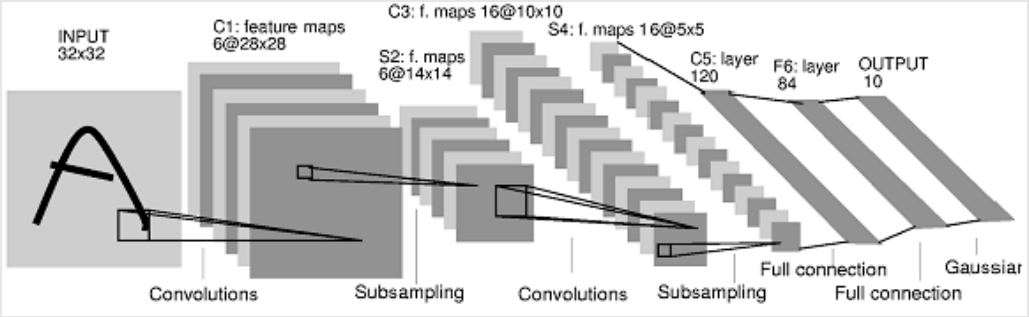
Outline

1. Historical Context
2. Structured Prediction & Deep CNNs
3. Deep Learning for Healthcare

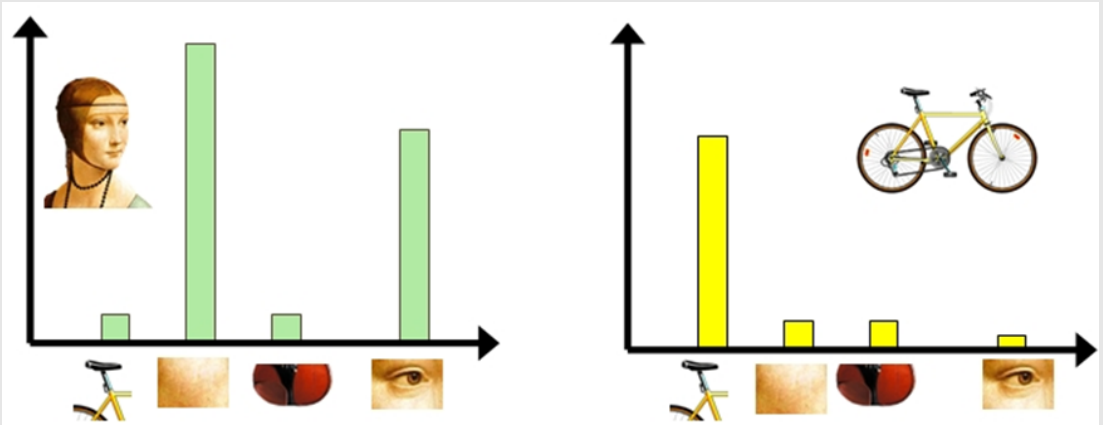
Visual Recognition History

Trends and methods in the last four decades

- 80's: training Convolutionnal Neural Networks (CNN) with back-propagation ⇒ postal code reading [LBD+89]



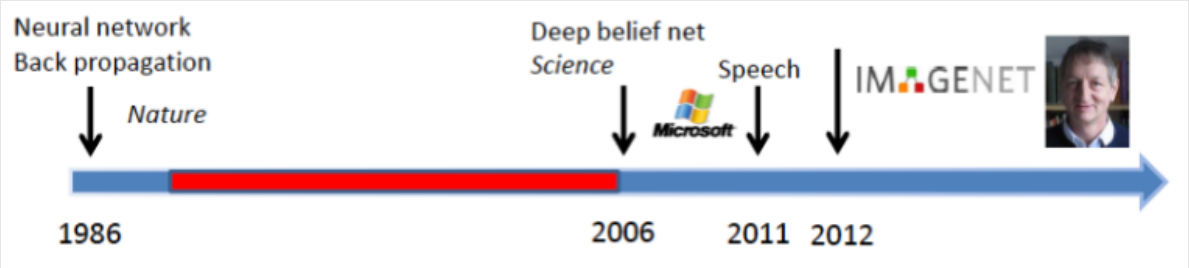
- 90's: golden age of kernel methods, NN = black box
- 2000's: BoW + SVM : state-of-the-art CV



Visual Recognition History

Trends and methods in the last four decades

- Deep learning revival: unsupervised learning (DBN) [HOT06]



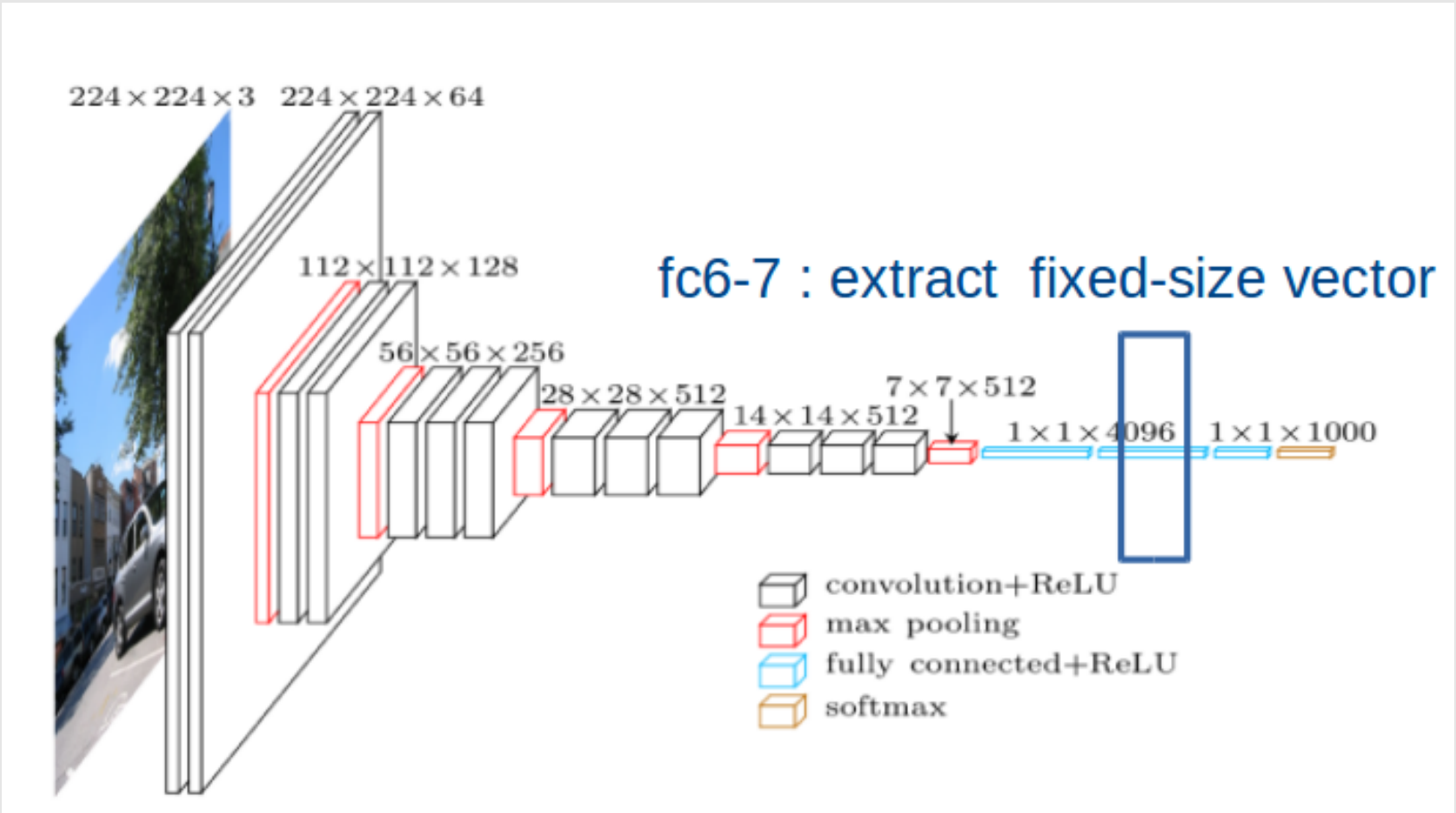
- 2012: CNN outstanding success in ImageNet [KSH12]

Rank	Name	Error rate	Description
1	**U. Toronto**	0.15315	Deep learning
2	U. Tokyo	0.26172	Hand-crafted features and learning models. Bottleneck.
3	U. Oxford	0.26979	
4	Xerox/INRIA	0.27058	

  - Huge number of labeled images (10^6 images)
  - GPU implementation for training

Deep Learning since 2012

Transferring Representations learned from ImageNet



- Extract layer ⇒ fixed-size vector, "deep features"
- Now state-of-the for any visual recognition task

Outline

1 Historical Context

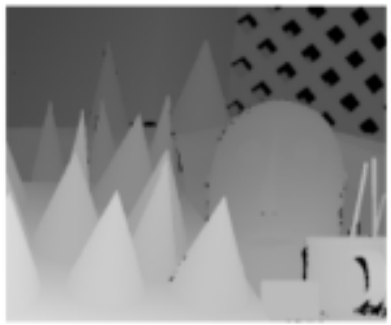
2 Structured Prediction & Deep CNNs

3 Deep Learning for Healthcare

Structured prediction and Deep CNNs

Deformations in Deep CNNs

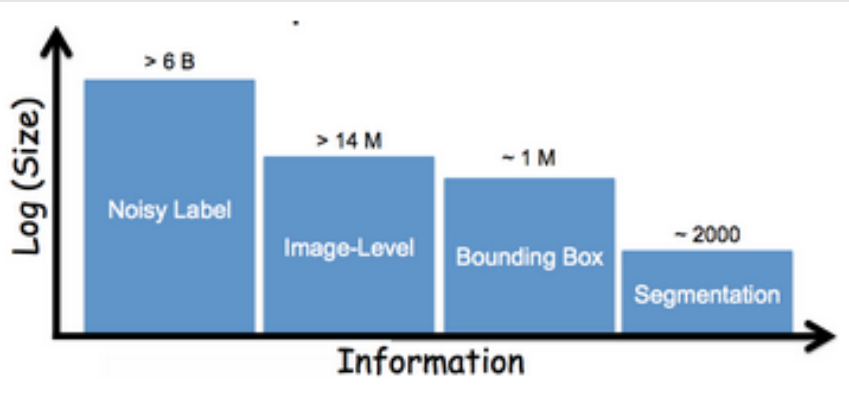
- Current CNN: only handle limited geometrical deformations (pooling layers)
- Strong deformations not dealt with
- Probabilistic graphical models: sound for structure modeling
  - Conditional Random Filed (CRF) & Structural SVM (SSVM)
- Crucial in various visual recognition tasks: detection, segmentation, pose estimation, depth estimation, sequence labeling, *etc*

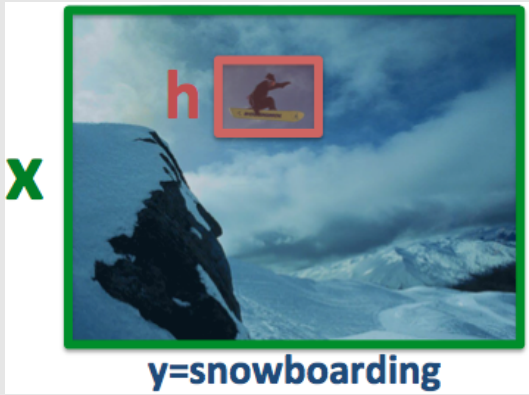



Structured prediction and Deep CNNs

Weakly Supervised Learning

- Detailed annotations: costly and tedious to collect
  - *e.g.* # global image labels >> # segmentation masks
- Option : using coarse annotations, *e.g.* global image labels
  - Incorporating latent variables (region selection)
  - Models : LSSVM, HCRF





Weakly Supervised Structured Deep Prediction

MANTRA [DTC15]: new structured output latent variable model

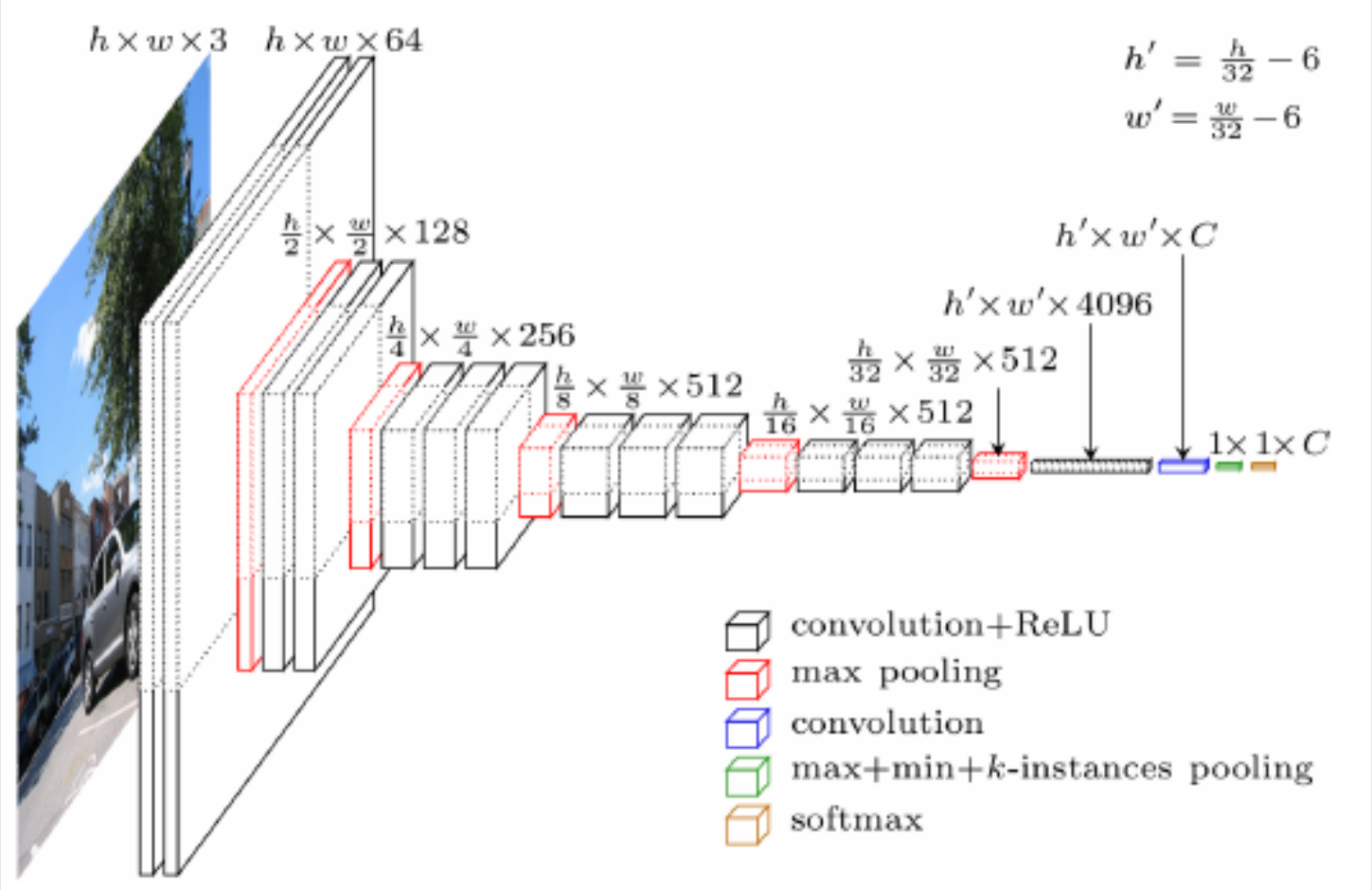
- Minimum Maximum Latent Structural SVM for Image Classification and Ranking
- New region selection strategy: max + min pooling
  - max: indicator of the **presence** of the class
  - min: indicator of the **absence** of the class

street image | **street** model | **coast** model | **highway** model

Weakly Supervised Structured Deep Prediction: WELDON

Weakly Supervised Learning of Deep CNNs [DTC16]

- MANTRA extension for end-to-end training of Deep CNNs
- Multiple max and min regions



Outline

1 Historical Context

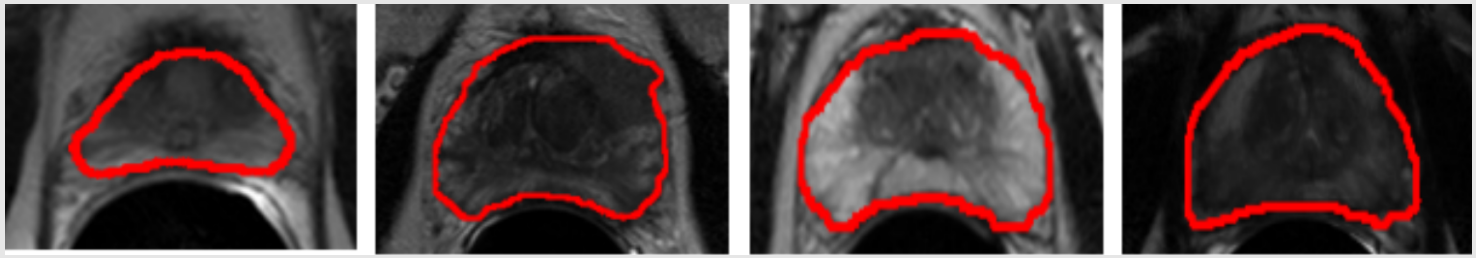
2 Structured Prediction & Deep CNNs

3 Deep Learning for Healthcare

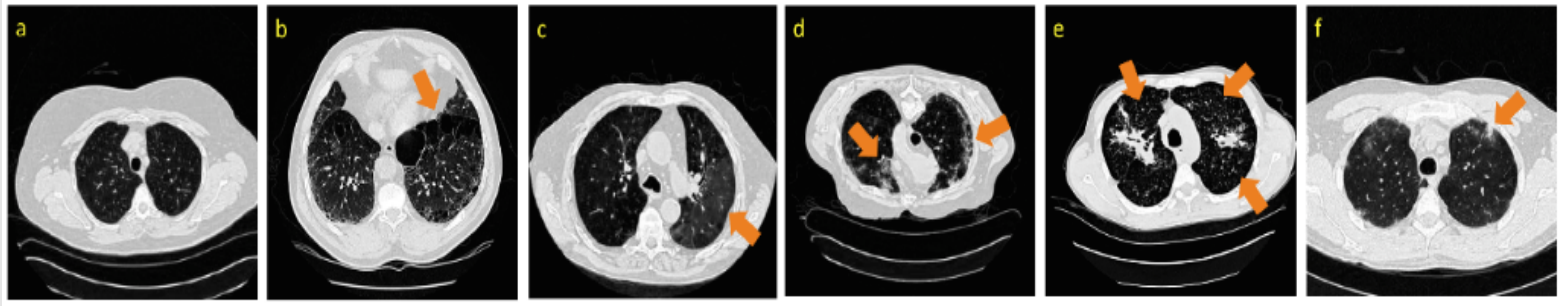
Deep learning for Healthcare

Medical image/video analysis

- Using deep learning for medical image annotation
  - Unsupervised representation learning for disease diagnosis, *e.g.* Alzeihmer [SS13], Prostate labeling [LGOS13]



  - Supervised learning with CNNs [SRG+16, TSG+16]

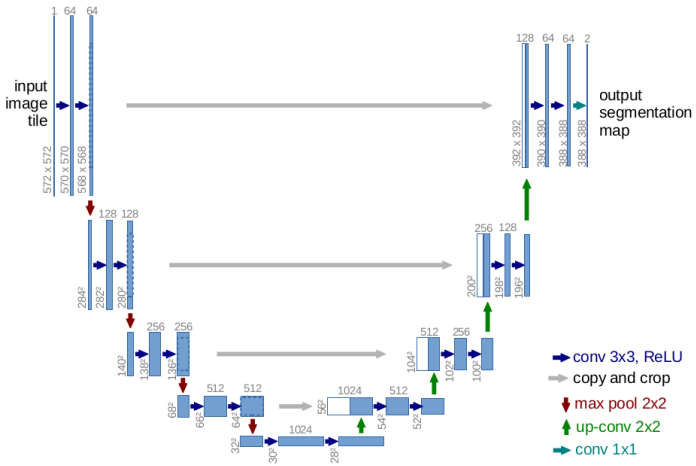


  - Transfer from ImageNet, fine-tuning, training from scratch ?
  - Impact of network size ?

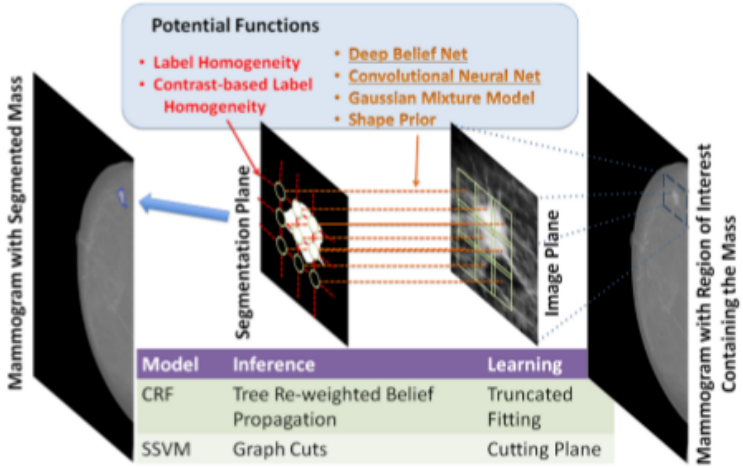
Deep learning for Healthcare

Medical image/video analysis: structured prediction with CNNs

- Using specific network architecture, *e.g.* U-Net [RFB15]
- Importance of structure ⇒ Embed CNN as a part of the overall formulation (*e.g.* CRF, SSVM)
  - Registration, segmentation [DCB15], weakly-supervised context [CPBN15]



Deep Convolutional Encoder Networks for Multiple Sclerosis Lesion Segmentation [RFB15]



Deep Learning and Structured Prediction for the Segmentation of Mass in Mammograms [DCB15]

Deep learning for Healthcare

Classification of Surgery Videos: APTITUDE Project (LIP6)

- Assistance à l'aPprenTIssage des gesTes chirUrgicaux à partir de viDEos ⇒ Application to cataract surgery
- Identify & segment different surgery steps, quantify its quality

- Deep CNN ⇒ surgery step
- *Prior* on step sequence ⇒ temporal smoothing (Vitterbi)

Thank you for your attention ! nicolas.thome@lip6.fr

Deep Learning for Visual Recognition

- Sorbonne Universités - UPMC Paris 6 - LIP6, lab, DAPA Dpt
  - Machine Learning and Information Access Team (P. Gallinari)
  - Machine Learning for Vision: N. Thome, M. Cord, ~ 10 PhD stud. http://webia.lip6.fr/~thomen/
- Weakly Supervised Learning of Deep CNNs [DTC15, DTC16]

  - Project Page: http://webia.lip6.fr/~durandt/project/mantra.html
- Surgery Video Classification, APTITUDE Project: L. Denoyer http://www-connex.lip6.fr/~denoyer/wordpress/

[DTC15] T. Durand, N. Thome, and M. Cord, MANTRA: Minimum Maximum Latent Structural SVM for Image Classification and Ranking, ICCV 2015.
[DTC16] T. Durand, N. Thome, and M. Cord, WELDON: Weakly Supervised Learning of Deep Convolutional Neural Networks, CVPR 2016.

References I

G. Carneiro, T. Peng, C. Bayer, and N. Navab, *Weakly-supervised structured output learning with flexible and latent graphs using high-order loss functions*, 2015 IEEE International Conference on Computer Vision (ICCV), Dec 2015, pp. 648–656.

Neeraj Dhungel, Gustavo Carneiro, and Andrew P. Bradley, *Medical image computing and computer-assisted intervention – miccai 2015: 18th international conference, munich, germany, october 5-9, 2015, proceedings, part i*, ch. Deep Learning and Structured Prediction for the Segmentation of Mass in Mammograms, pp. 605–612, Springer International Publishing, Cham, 2015.

Thibaut Durand, Nicolas Thome, and Matthieu Cord, *MANTRA: Minimum Maximum Latent Structural SVM for Image Classification and Ranking*, International Conference on Computer Vision (ICCV), 2015.

Thibaut Durand, Nicolas Thome, and Matthieu Cord, *WELDON: Weakly Supervised Learning of Deep Convolutional Neural Networks*, Computer Vision and Pattern Recognition (CVPR), 2016.

Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh, *A fast learning algorithm for deep belief nets*, Neural Comput. 18 (2006), no. 7, 1527–1554.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, *Imagenet classification with deep convolutional neural networks*, Advances in neural information processing systems, 2012, pp. 1097–1105.

Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel, *Backpropagation applied to handwritten zip code recognition*, Neural computation 1 (1989), no. 4, 541–551.

References II


Shu Liao, Yaozong Gao, Aytekin Oto, and Dinggang Shen, *Medical image computing and computer-assisted intervention – miccai 2013: 16th international conference, nagoya, japan, september 22-26, 2013, proceedings, part ii*, ch. Representation Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentation, pp. 254–261, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.


Olaf Ronneberger, Philipp Fischer, and Thomas Brox, *Medical image computing and computer-assisted intervention – miccai 2015: 18th international conference, munich, germany, october 5-9, 2015, proceedings, part iii*, ch. U-Net: Convolutional Networks for Biomedical Image Segmentation, pp. 234–241, Springer International Publishing, Cham, 2015.


H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, *Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning*, IEEE Transactions on Medical Imaging PP (2016), no. 99, 1–1.


Heung-Il Suk and Dinggang Shen, *Medical image computing and computer-assisted intervention – miccai 2013: 16th international conference, nagoya, japan, september 22-26, 2013, proceedings, part ii*, ch. Deep Learning-Based Feature Representation for AD/MCI Classification, pp. 583–590, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.


N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, *Convolutional neural networks for medical image analysis: Fine tuning or full training?*, IEEE Transactions on Medical Imaging PP (2016), no. 99, 1–1.