

Modélisation et Reconnaissance des Formes

Introduction

Marie-Odile Berger

<http://members.loria.fr/moberger>

September 13, 2023

Objectifs du cours

- ▶ cadre général: la reconnaissance et l'identification des formes
- ▶ **forme**= image nuage de points, signal de parole, des capteurs...
- ▶ Quelques problèmes:
 - ▶ reconnaître la présence d'un objet/des objets dans une image (mais pas seulement)
 - ▶ localiser ces objets dans l'image
 - ▶ connaître précisément sa position géométrique (ex: robotique, conduite automatisée...)
 - ▶ tenir compte de l'incertain sur l'identification pour effectuer une action

But

Modéliser des formes, les reconnaître, les localiser...

en tenant compte de l'incertain, des connaissances a priori, des données (exemples) disponibles

Objectif: Doter les machines des capacités de l'homme à reconnaître des caractères, des objets, des sons, des signes des signaux temporels...

Deux grands objets d'étude:

- ▶ Étudier de quelle manière l'être humain effectue cette reconnaissance (touche à des domaines comme psychologie, physiologie, biologie)
- ▶ Viser le développement de théories et de techniques permettant d'effectuer certaines tâches de reconnaissance (domaines: informatique, statistique, mathématiques)

Au départ, la RF est surtout du traitement du signal

- ▶ test de la présence d'un signal
- ▶ identification de sources multiples
- ▶ traitement de la parole

et progressivement, on a envisagé des tâches plus complexes...

L'humain fait beaucoup de choses:

reconnaître

- ▶ des visages
- ▶ des sons
- ▶ des formes

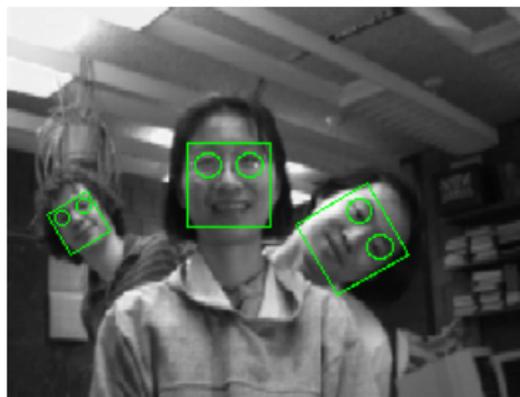
et ceci indépendamment

- ▶ du point de vue sous lesquels on les observe
- ▶ des conditions d'observation
- ▶ de leur variabilité

Le domaine est en forte progression depuis 2012 et l'émergence des réseaux convolutionnels.

- ▶ c'est par exemple
 - ▶ Identifier la présence d'objets dans une image
 - ▶ localiser l'objet dans l'image
 - ▶ le segmenter finement
- ▶ mais aussi...
 - ▶ identifier des structures (souvent paramétriques) dans des données (images, nuages de points, mesures industrielles...)
 - ▶ estimer l'état d'un modèle à partir d'observations directes ou indirectes
 - ▶ construire des modèles de formes
 - ▶ valider l'appartenance d'une donnée à un modèle
 - ▶ ...

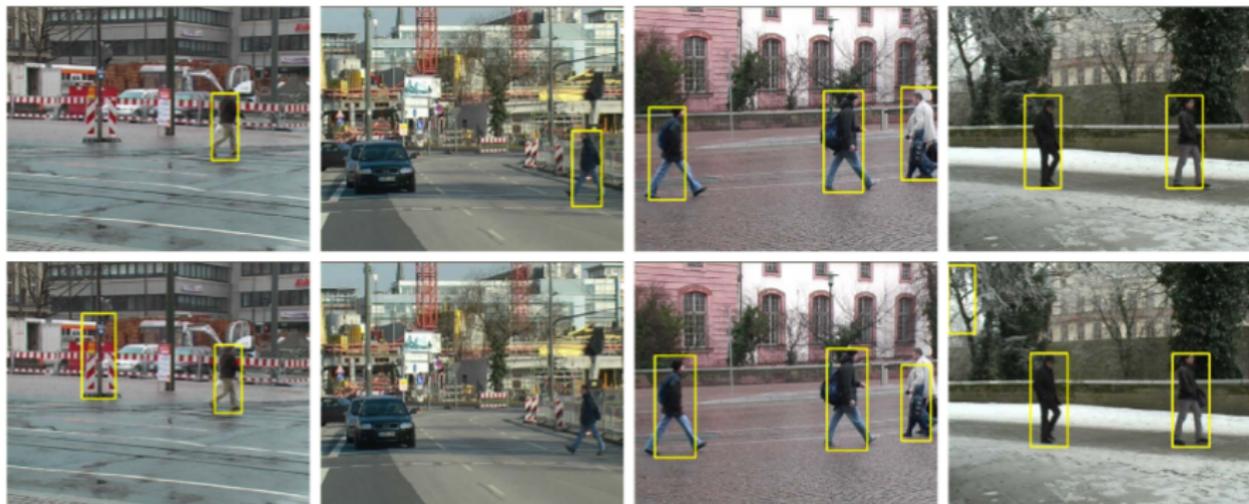
Quelques exemples déjà anciens en images



Détection précise ou approximative?

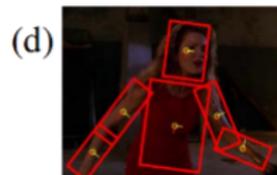
Détection = localisation des objets dans l'image

Détection de piétons [Andriluka CVPR 2009]: on peut se contenter de détecter des boîtes contenant des piétons



Détection précise ou approximative?

Ou avoir une estimation précise de la position du corps



Segmentation sémantique

but: attribuer à chaque pixel une classe d'appartenance (sol, arbre, voiture,...)

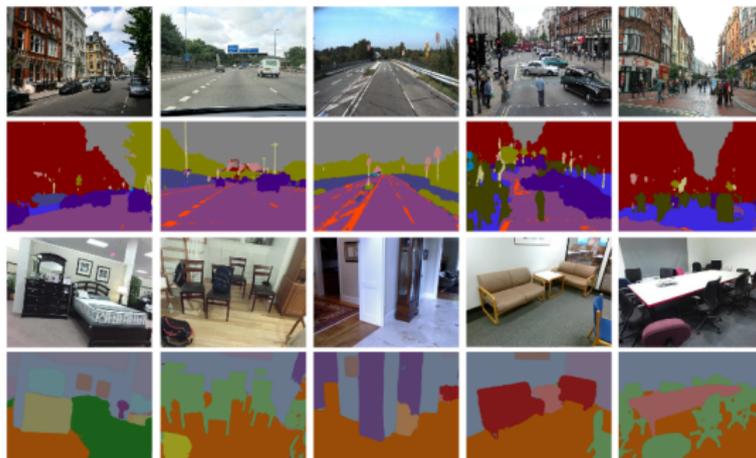
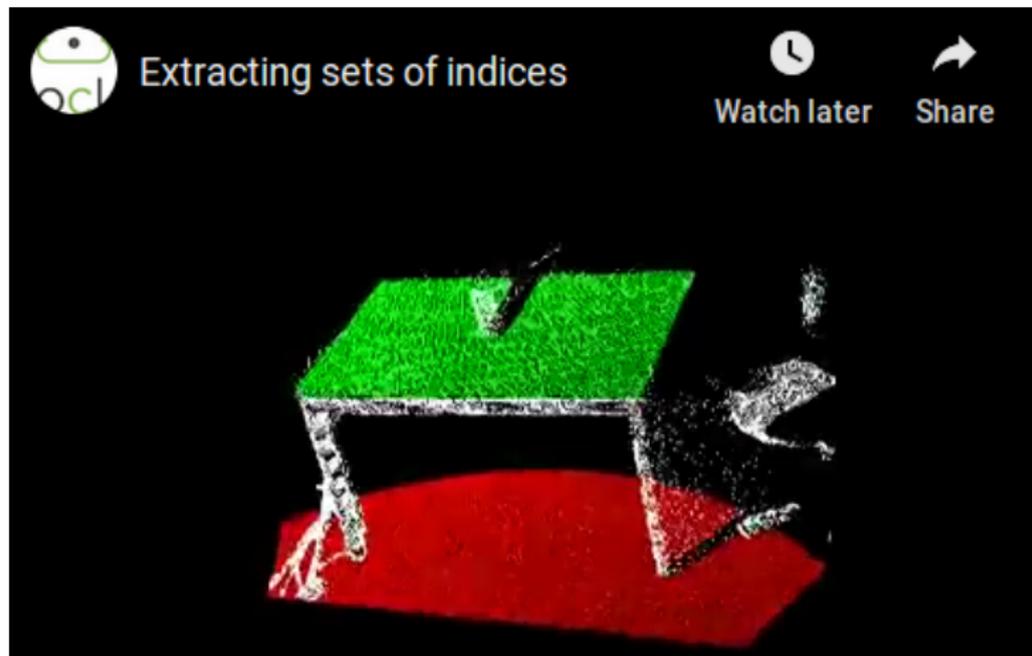


Figure: Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Badrinarayanan 2016]

Exemple: extraction de plans dans un nuage de points



Ici, l'objet cherché a un modèle paramétrique

Joint reconstruction and semantic segmentation [Hane CVPR 2013]

Où comment la résolution conjointe de deux problèmes améliore la reconstruction → régularisation spécifique à chaque classe

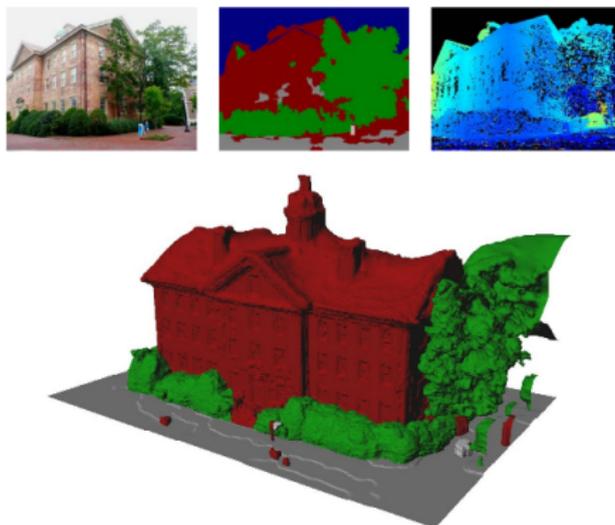


Figure 1: Top: Example of input image, standard image classification result, depthmap. Bottom: Our proposed joint optimization combines class segmentation and geometry resulting in an accurately labeled 3D reconstruction

Cognitivistes contre comportementalistes

- ▶ le cognitiviste s'inspire de nos connaissances sur la perception humaine pour concevoir des systèmes d'inspiration cybernétique.
- ▶ le comportementaliste ne cherche pas à analyser le concept mais essaie de collecter un maximum de données différentes pour en extraire des régularités et des moyens de classification

→ avec l'expansion des CNNs, c'est l'approche travaillant directement sur les données qui a pris l'ascendant.

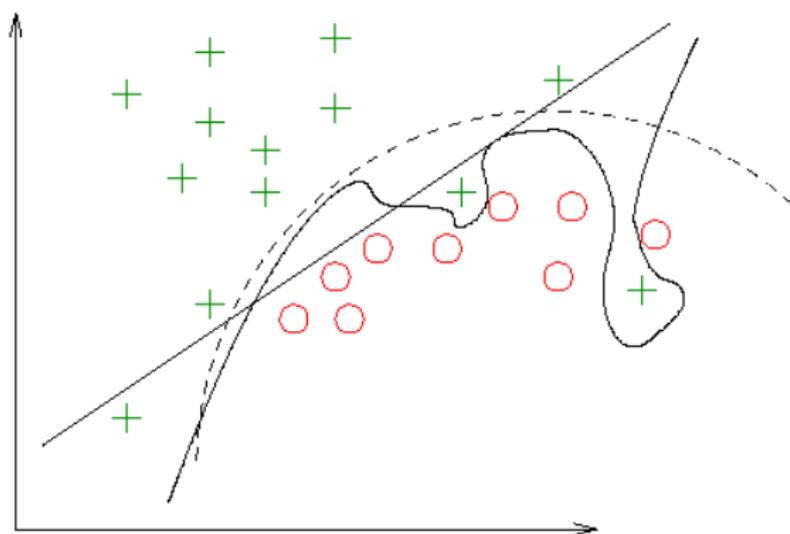
Exemple de l'évolution de la reconnaissance de visages:

- ▶ on est passé d'approches visant à extraire yeux, bouches, lèvres (représentation explicite)
- ▶ à des approches où on apprend une représentation linéaire d'après une base de données (voir le cours sur l'ACP et le TP sur les eigenfaces [11])
- ▶ puis à des approches avec des CNNs (DeepFace [10] et poster **poster** associé) où on donne des millions d'exemples et le système construit une représentation interne des visages (exemple base de données d'Oxford, dont la dernière <https://www.robots.ox.ac.uk/vgg/data/voxceleb/>)
- ▶ **Précautions d'usage:** Comment savoir si les données recueillies sont bien représentatives du phénomène? Consulter l'article 2018 de **Sciences et Avenir**: Intelligence artificielle : la reconnaissance faciale est-elle misogyne et raciste ?

Voir aussi **Une brève histoire de la reconnaissance de visages** avec les avancées technologiques et usages récents

En présence de données aberrantes ou atypiques?

Le contrôle d'une forme de régularité des classes ou des distributions, permet d'éviter de créer des frontières inutilement complexes, risquant de refléter uniquement le bruit...



Gérer les données incorrectes est un problème important de la RF.

Problèmes typiques de la RF

Une série de problèmes typiques de la RF que nous aborderons via les **méthodes conventionnelles** ou les **réseaux convolutionnels**.

- ▶ **Comparer les formes**: réduction de la dimensionalité, mesure de similarité, invariance au contexte...
- ▶ **Modéliser/représenter les formes**: par des méthodes explicites ou des réseaux de neurones
- ▶ **robustesse**: les résultats des algorithmes doivent (devraient) être insensibles à la présence de données aberrantes (régularisation, estimation robuste)
- ▶ **estimation**: calculer les paramètres d'un modèle à partir de données (problème d'optimisation car la plupart des problèmes modélisés par le minimum d'une fonction de coût)

- ▶ Représenter et comparer les formes: des descripteurs construits à la main aux descripteurs CNN
- ▶ Données et incertitudes
 - ▶ Modéliser et propager l'incertain
- ▶ Modélisation: le cas linéaire
 - ▶ Régression linéaire: moindres carrés
 - ▶ l'analyse en composante principales (ACP)
- ▶ Modéliser et regresser avec des réseaux de neurones
 - ▶ MLP, réseaux siamois, réseaux adversoriaux (GAN)
- ▶ Estimation et optimisation
 - ▶ Les méthodes classiques d'estimation: moindres carrés, max vraisemblance, ...
 - ▶ Méthodes classiques d'optimisation (descente de gradient) , Algo EM
 - ▶ Estimation robuste



S. Agarwal and D. Roth.

Learning a sparse representation for object detection.

In *Proceedings of 7th European Conference on Computer Vision, Copenhagen (Denmark)*, 2002.



Mykhaylo Andriluka, Stefan Roth, and Bernt Schiele.

Pictorial structures revisited: People detection and articulated pose estimation.

In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2009.

Best Paper Award Honorable Mention by IGD.



Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla.

Segnet: A deep convolutional encoder-decoder architecture for image segmentation.

CoRR, abs/1511.00561, 2015.

Bibliographie II

 T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham.

Active shape models -their training and application.

Computer Vision and Image Understanding, 61(1):38–59, 1995.

 R. O. Duda and P. E. Hart.

Pattern Classification and Scene Analysis.

Wiley-InterScience, 1973.

 C. Hane, C. Zach, A. Cohen, R. Angst, and M. Pollefeys.

Joint 3d scene reconstruction and class segmentation.

In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 97–104, June 2013.

 A. K. Jain, R. P. W. Duin, and J. Mao.

Statistical Pattern Recognition: A Review.

IEEE Transactions on PAMI, 22(1):4–37, January 2000.

Bibliographie III



Yann LeCun, Fu-Jie Huang, and Leon Bottou.

Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting.

In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC (USA), 2004.



L. Rabiner.

A tutorial on hidden markov models and selected applications in speech recognition.

Proc. IEEE, 77:257–286, 1989.



Y. Taigman, M. Yang, M. Ranzato, and L. Wolf.

Deepface: Closing the gap to human-level performance in face verification.

In 2014 IEEE Conference on Computer Vision and Pattern Recognition, pages 1701–1708, June 2014.



M. Turk and A. Pentland.

Eigenfaces for Recognition.

Journal of Cognitive Neuroscience Volume 3, Number 1, 3(1):71–86, 1991.



A. Webb, editor.

Statistical Pattern Recognition.

wiley, 2002.