

# Détection d'arbres urbains à partir de données aériennes multi-sources (optique, infrarouge, MNS)

Lionel PIBRE<sup>1,5</sup>, Dino IENCO<sup>1,4</sup>, Gérard SUBSOL<sup>1,3</sup>  
Mustapha DERRAS<sup>5</sup>, Marc CHAUMONT<sup>1,2,3</sup>

<sup>1</sup>Université de Montpellier

<sup>2</sup>Université de Nîmes

<sup>3</sup>CNRS

<sup>4</sup>IRSTEA

<sup>5</sup>Berger-Levrault

Atelier Télédétection pour l'Etude des Milieux Urbains 2017  
16 Janvier 2017

- 1 Introduction
- 2 Notre approche
- 3 Résultats
- 4 Perspectives

# Problématique

Dans notre cas, nous sommes sur un problème de détection  $\neq$  segmentation.

- Détection  $\Rightarrow$  Approche objet
- Segmentation  $\Rightarrow$  Approche pixel

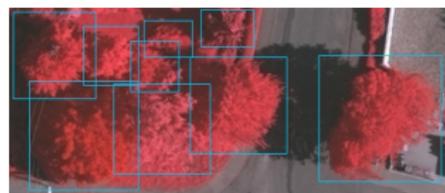


FIGURE: Exemple de détection.



FIGURE: Exemple de segmentation.

# La segmentation

- La segmentation donne de bons résultats  $\Rightarrow$  Passage à l'objet non trivial,
- L'approche pixel est souvent insuffisante pour dissocier les arbres,
- On veut localiser tous les individus  $\Rightarrow$  Prise en compte globale de l'objet.

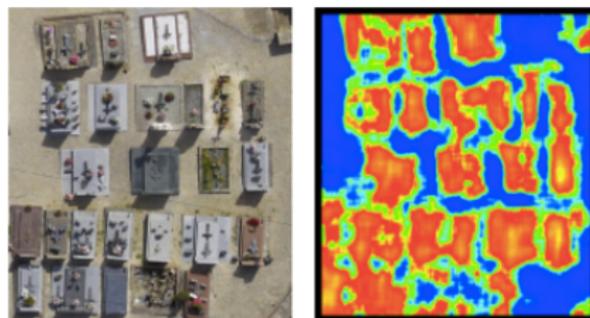


FIGURE: Exemple de segmentation sur des tombes.

# La localisation

Afin de localiser les arbres dans les images, nous avons utilisé une fenêtre glissante multirésolution

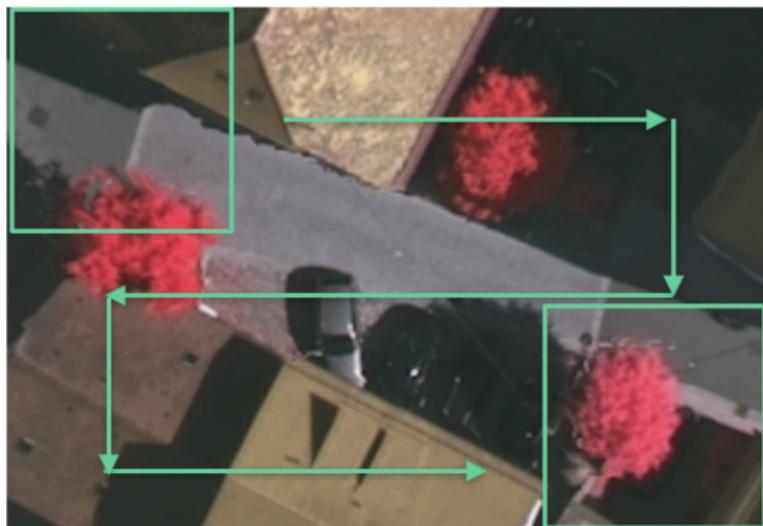


FIGURE: Application de la fenêtre glissante sur une image.

# Quelle méthode utiliser ?

- La méthode classique consiste à extraire des caractéristiques pour chaque type de données et ensuite faire une classification.
- Exemple de descripteurs : HOG, SIFT ...
- Exemple de classifieurs : SVM, Random Forests ...
- Plus récent : utilisation d'un réseau de neurones convolutionnels (CNN).

- 1 Introduction
- 2 Notre approche**
- 3 Résultats
- 4 Perspectives

Comparaison entre deux CNNs et deux méthodes de machine learning sur la base de données Vaihingen.

Comparaison selon le rappel, la précision et la F-mesure entre :

- AlexNet
- GoogleNet
- HOG + SVM
- HOG + Random Forests

# la méthode utilisée

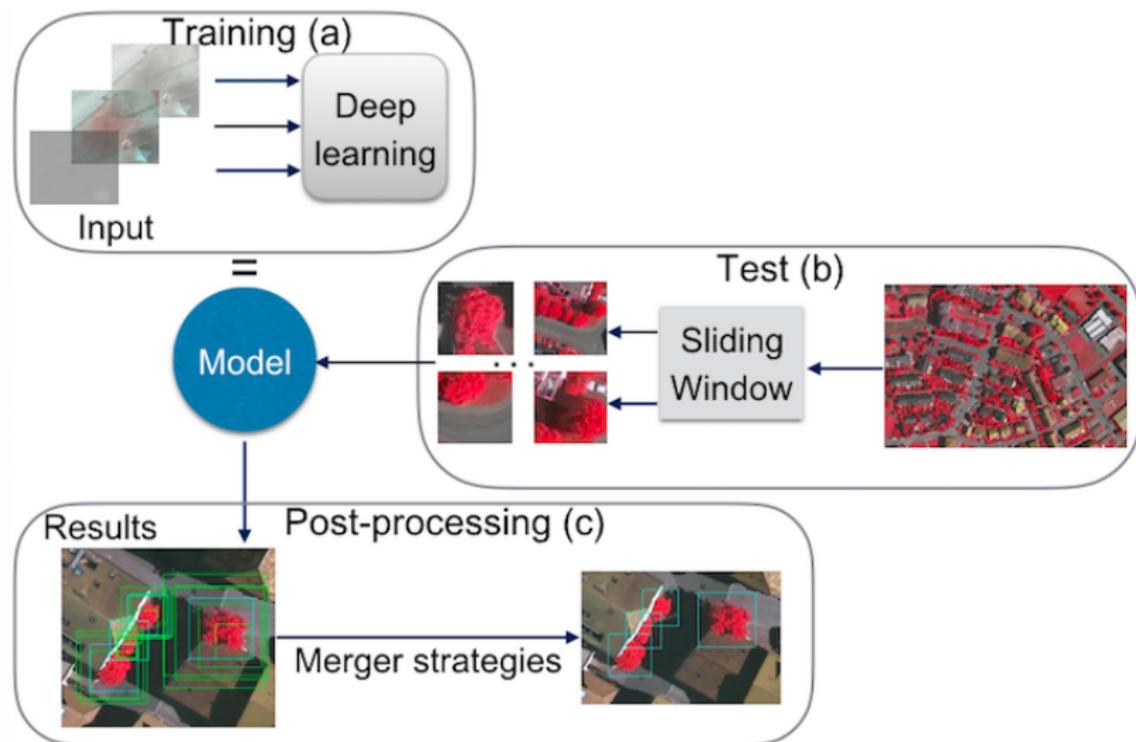
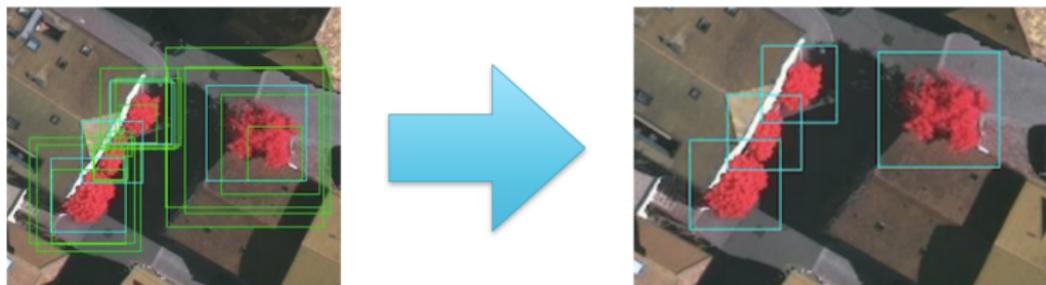


FIGURE: Schéma de la méthode proposée.

# Fusion des boîtes englobantes

Réduction du nombre de boîtes englobantes multiples et donc de faux positifs.



**FIGURE:** Fusion des boîtes englobantes.

Caractéristiques de cette base de données :

- Résolution au sol des images R,G,PIR : 8cm
- Résolution au sol du MNS : 9cm
- Acquisition sur 3 zones :
  - Centre ville
  - Zone résidentielle
  - Quartier d'affaires
- Acquisition des images optiques : Intergraph / ZI DMC
- Acquisition du MNS : Leica ALS50

# Les images de Vaihingen

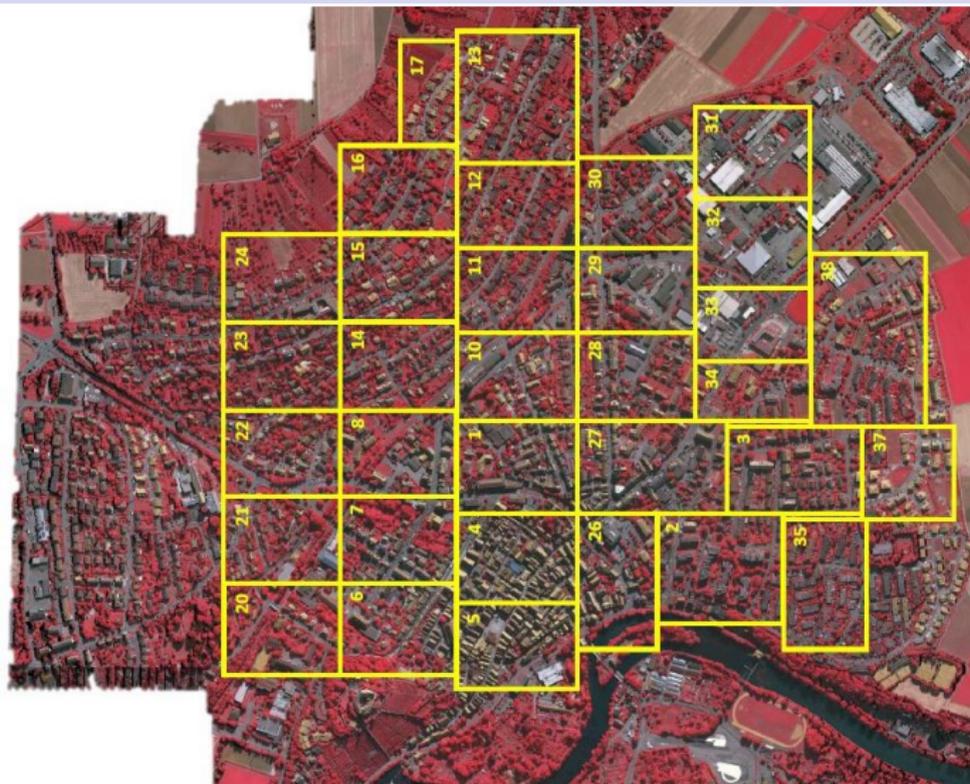


FIGURE: Zones d'acquisition de la base de données Vaihingen.

- 1 Introduction
- 2 Notre approche
- 3 Résultats**
- 4 Perspectives

# Les mesures utilisées

Afin de comparais les différentes méthodes, nous avons utilisé les mesures suivantes :

$$Rappel = \frac{TP}{TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Fmesure = \frac{2RP}{P + R} \quad (3)$$

$$label = \begin{cases} 1 & \text{If } \frac{area(detection \cap ground\ truth)}{area(detection \cup ground\ truth)} > 0.5 \\ 0 & \text{If } \frac{area(detection \cap ground\ truth)}{area(detection \cup ground\ truth)} \leq 0.5 \end{cases} \quad (4)$$

Avec  $TP$  les vrais positifs,  $FP$  les faux positifs et  $FN$  les faux négatifs.

TABLE: Résultats des différentes méthodes sur la base de données Vaihingen.

	AlexNet	GoogleNet	HOG+SVM	HOG+RF
Aire				
Rappel	41.46%	<b>46.99%</b>	26.66%	38.67%
Précision	24.28%	<b>29.24%</b>	0.95%	7.77%
F-Mesure	30.41%	<b>35.68%</b>	1.83%	10.91%
Chevauchement				
Rappel	<b>49.28%</b>	48.96%	21%	33.47%
Précision	22.57%	<b>25.71%</b>	1.54%	10.47%
F-Mesure	30.63%	<b>33.32%</b>	2.88%	13.78%

- Le descripteur HOG seul n'est pas efficace pour cette tâche,
- Les CNNs donnent de meilleures performances,
- La fusion par aire est plus efficace sur les CNNs,
- Cette tâche est difficile :
  - Les arbres sont collés,
  - De nombreux objets ont la même apparence.

- 1 Introduction
- 2 Notre approche
- 3 Résultats
- 4 Perspectives**

- Faire du transfer learning,
- Utilisation du NDVI,
- Ajouter du pré-traitement :
  - Utiliser le NDVI et le MNS pour ne pas traiter les zones sans végétation,
- Intégration de la localisation dans les réseaux de neurones convolutionnels,
- Combiner la segmentation et la détection.

*Merci de votre attention!*

