

Localisation et comptage d'arbres urbains dans des images multispectrales aériennes

19/10/2017

Lionel Pibre

Sous la direction de Marc Chaumont, Gérard Subsol, Dino Ienco et Laurent Deruelle



- **CONTEXTE**
- **BASE DE DONNÉES**
- **ARCHITECTURES**
- **RÉSULTATS**
- **LES AMÉLIORATIONS POSSIBLES**

Contexte

- Thèse CIFRE
 - Berger-Levrault
 - Laboratoire d'Informatique, Robotique et Microélectronique de Montpellier (LIRMM)
- Équipe ICAR
 - Images et Interactions
 - Spécialisée dans le traitement des images
- Encadrants au laboratoire
 - Marc Chaumont
 - Gérard Subsol
 - Dino Ienco
- Encadrant chez Berger-Levrault
 - Laurent Deruelle

Qu'est ce que l'on veut faire?

Localiser les arbres dans des images multispectrales aériennes + LiDAR

Quelle est la difficulté?

Les arbres sont collés, parfois cachés par d'autres objets urbains

Comment fusionner deux/plusieurs types de données?

Pourquoi faire?

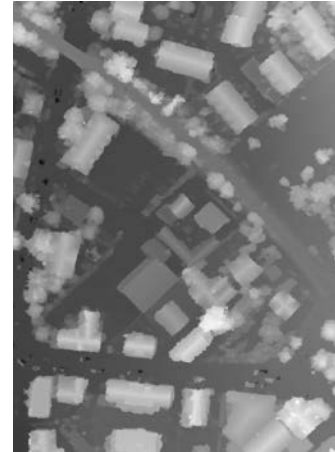
À court terme: les compter et les localiser

À long terme: Déterminer l'essence des arbres détectés et leur état de santé

Vaihingen¹

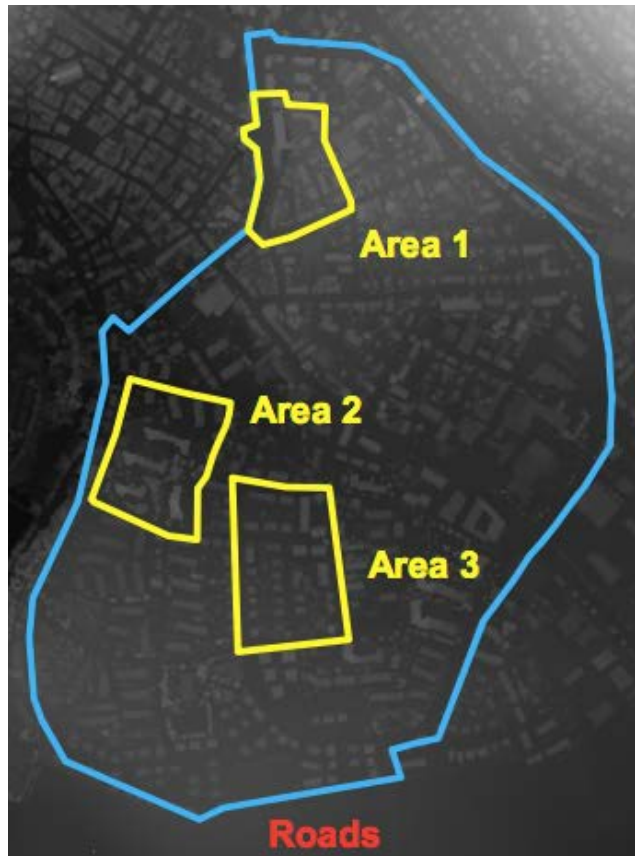
Caractéristiques de cette base de données :

- Résolution au sol des images R,V,PIR : 8 cm
- Résolution au sol du MNS : 9 cm
- L'acquisition contient 3 types de zones :
 - Centre ville
 - Zone résidentielle
 - Immeubles
- Acquisition des images optiques : Intergraph / ZI DMC
- Acquisition du MNS : Leica ALS50

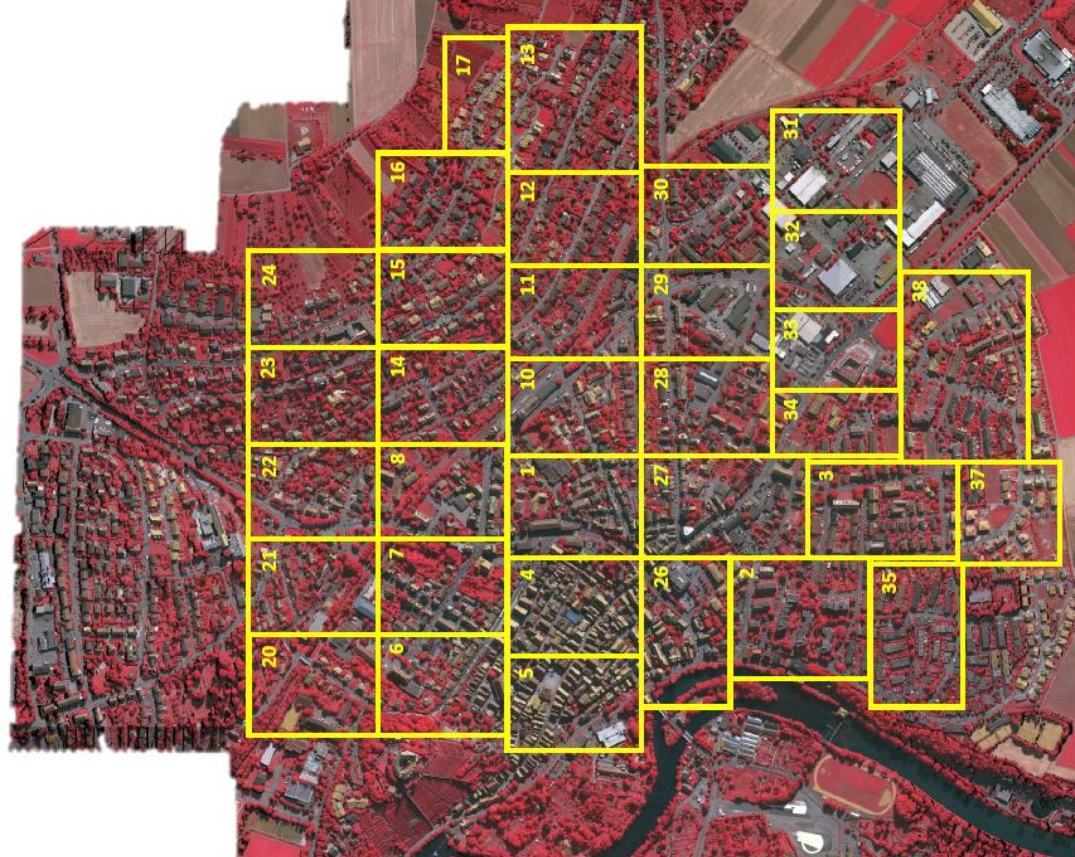


1. M. Cramer, "The dgpf-test on digital airborne camera evaluation—overview and test design," *Photogrammetrie-Fernerkundung-Geoinformation*, vol. 2010, no. 2, pp. 73–82, 2010.

Vaihingen



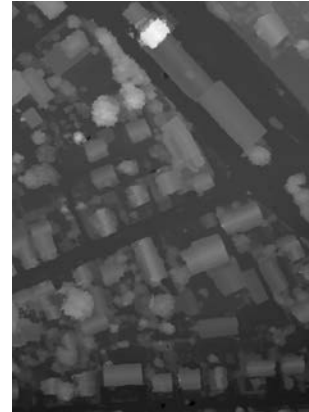
Vaihingen



- 38 images de taille variable
- Contenu hétérogène
 - ▶ Images 4 et 5: très peu de végétation
 - ▶ Images 20 et 24: végétation abondante
- Annotation pixel disponible
 - ▶ Végétation haute
 - ▶ Végétation basse
 - ▶ Bâtiment
 - ▶ Voiture
 - ▶ Route
 - ▶ Autre

Base de données

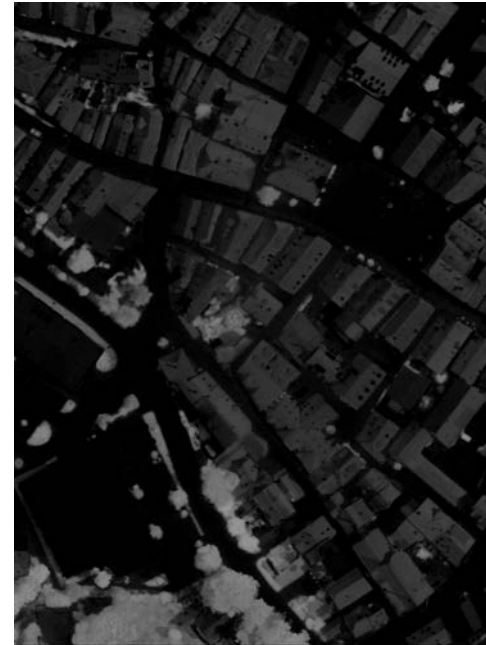
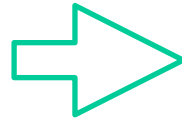
- Base de données:
 - ▶ Vaihingen
 - ▶ Images: R,V,PIR et MNS
- Base d'entraînement:
 - ▶ 6 000 images "arbre", 40 000 images "autre"
 - ▶ Taille: 64x64
 - ▶ Augmentation des images "arbres" en utilisant 3 rotations (1 500 images à l'origine)
- Base de test:
 - ▶ 20 images de taille variable (de 125x150 jusqu'à 550x775)
 - ▶ Une centaine d'arbres



Pré-traitement des données

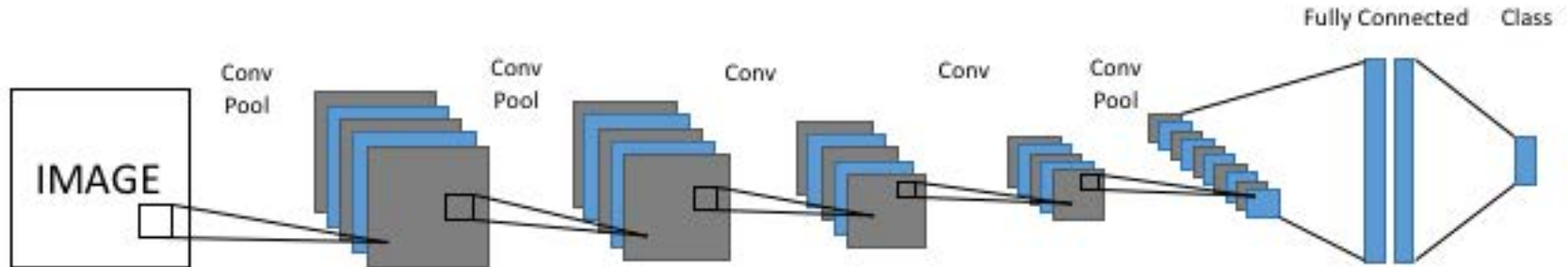
- Utilisation du NDVI

$$NDVI = \frac{PIR - R}{PIR + R}$$



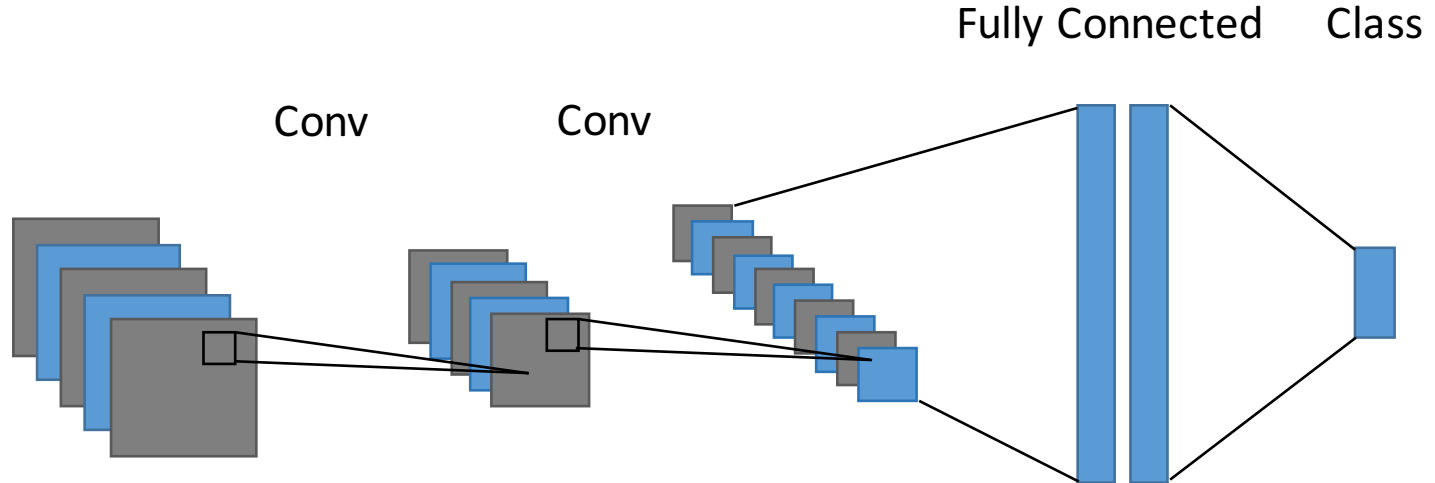
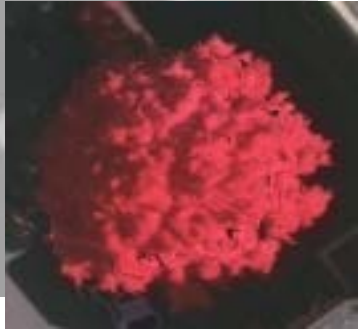
Architecture du réseau

- Réseau de base: AlexNet
- Deux stratégies de fusion ont été utilisées:
 - ▶ Early Fusion
 - ▶ Late Fusion



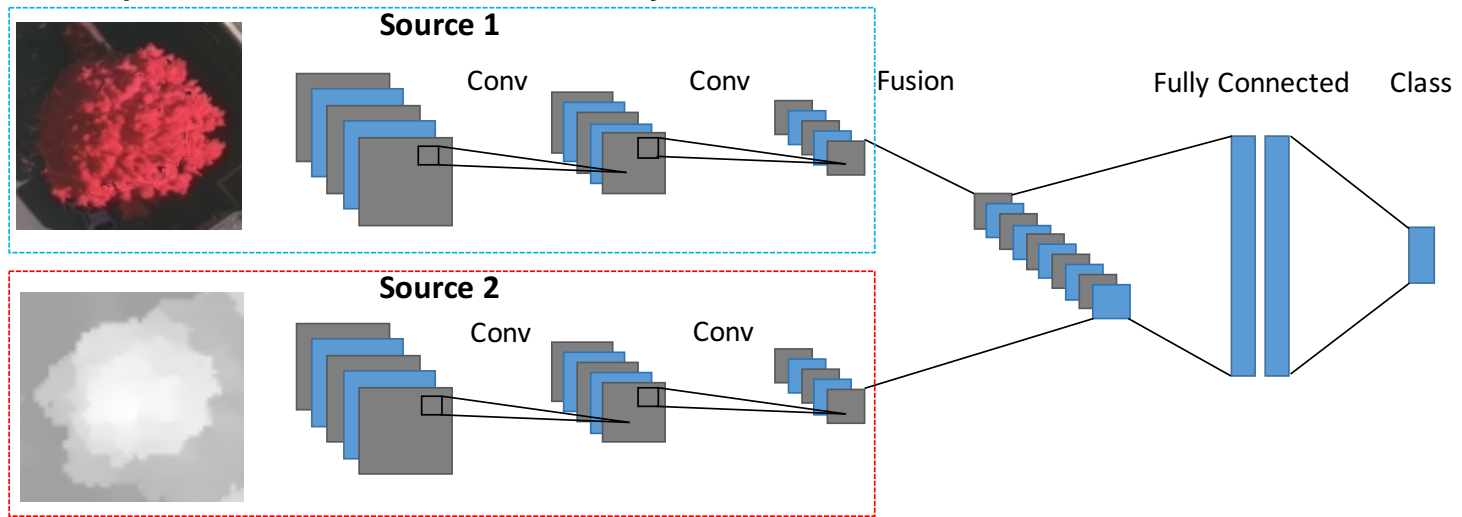
Early Fusion Architecture

- Traitement des différentes informations en même temps
- Trouver une optimisation optimale en combinant les informations



Late Fusion Architecture

- Traiter les différentes informations indépendamment
- Architecture d'AlexNet
- Traitement dans des branches séparées du réseau
- Concaténation juste avant les couches Fully Connected



Expériences

- Expériences réalisées avec une 5-fold cross validation
- Base d'entraînement:
 - ▶ 6 000 images “arbre”, 40 000 images “autre”
 - ▶ Taille: 64x64
 - ▶ Augmentation des images “arbres” en utilisant 3 rotations (1 500 images à l'origine)
- Base de test:
 - ▶ 20 images de taille variable (de 125x150 jusqu'à 550x775)
 - ▶ Une centaine d'arbres
- On choisit le point sur la courbe Rappel/Précision où la F-Mesure est la plus élevée

Expériences

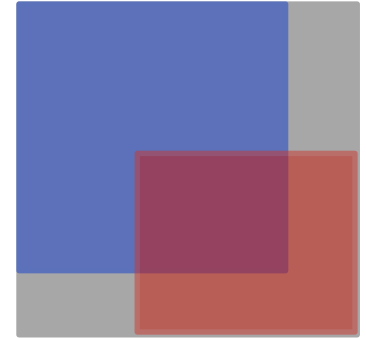
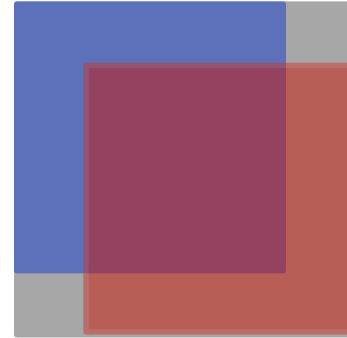
- Quelles évaluations ont été utilisées?

$$label = \begin{cases} 1 & \text{If } \frac{aire(detection \cap verite\ terrain)}{aire(detection \cup verite\ terrain)} > 0.5 \\ 0 & \text{If } \frac{aire(detection \cap verite\ terrain)}{aire(detection \cup verite\ terrain)} \leq 0.5 \end{cases}$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

$$F - Measure_{max} = 2 \frac{Recall * Precision}{Recall + Precision}$$



- TruePositives
- FalsePositives
- FalseNegatives



Limites d'utiliser une seule source

SOURCE	RVPIR	MNS	NDVI
F-Mesure _{max}	60,45 %	62,47 %	63,97 %
Rappel	57,89 %	57,62 %	62,34 %
Précision	63,44 %	68,56 %	67,04 %

- Les résultats des F-Mesure entre les différentes sources sont très proches (moins de 4% d'écart)
- La meilleure F-Mesure est obtenue en utilisant le NDVI
- On obtient la meilleure précision avec le MNS

Quelle est la meilleure méthode de fusion? - Early Fusion Architecture

Early Fusion	RVPIR+MNS	NDVI+MNS
F-Mesure _{max}	67,12 %	75,30 %
Rappel	65,40 %	68,37 %
Précision	69,54 %	84,11 %

- La meilleure F-Mesure est obtenue en utilisant le NDVI+MNS
- Le rappel entre les deux tests est très proche, seulement 3% de différence
- Le gain apporté par le NDVI sur la Précision est de presque 15%

Quelle est la meilleure méthode de fusion? - Late Fusion Architecture

Late Fusion	RVPIR+MNS	NDVI+MNS
F-Mesure _{max}	62,14 %	72,57 %
Rappel	62,54 %	70,99 %
Précision	62,65 %	74,83 %

- La meilleure F-Mesure est obtenue en utilisant le NDVI+MNS
- Le rappel entre les deux tests est plus important, plus de 8% de différence
- Le gain apporté par le NDVI sur la Précision n'est que de 12%

Comment sélectionner les sources à fusionner?

SOURCES	RVPIR/MNS		NDVI/MNS	
Corrélation	47,86 %		48,96 %	
Distribution	26,47 %	25,66 %	28,75 %	22,27 %

- Environ 50% de tous les arbres trouvés ont été trouvés dans les 2 sources
- Les 50% restant sont répartis sur les 2 sources
 - ▶ Cela montre l'intérêt d'utiliser plusieurs sources
- La corrélation des faux positifs entre les 2 sources a elle aussi été calculée mais elle ne dépasse pas les 10%
 - ▶ La combinaison des sources devrait réduire le nombre de faux positifs

Résultats

- L'architecture Early Fusion permet d'obtenir une meilleure F-Mesure (75% contre 72,5%)
- Les meilleurs résultats sont obtenus lorsque l'on fusionne le NDVI avec le DSM
 - ▶ Le NDVI permet de ne garder que ce qui est de la végétation
 - ▶ Le DSM nous donne la hauteur à toutes les positions dans l'image
 - ▶ Si un objet est en hauteur et est de la végétation
 - ✓ Il y a de fortes chances pour que ce soit un arbre

Améliorations possibles

- Faire du transfer learning pour utiliser un modèle pré-entraîné
 - ▶ Utiliser la base de données ImageNet par exemple
- Utiliser une autre architecture
 - ▶ VGG
 - ▶ ResNet
- Intégrer la détection dans le réseau
 - ▶ Fast-RCNN

Fin



Merci

“Il est facile de poser la question difficile.”

W. H. Auden