

# Identification / Localisation de fruits dans des images couleur d'arbres : application au manguier

Philippe Borianne (Amap)

[philippe.borianne@cirad.fr](mailto:philippe.borianne@cirad.fr)

Restitution stage M2 SNS – parcours BCD

Guilhem BARTHÈS

*22 mai – 11 août 2017*

Action Incitative 2017 Créativité et innovation scientifique CRESI

### **Projet PixYield – É. Faye (UR HortSys)**

*Analyse d'image & Estimation de rendement*



# un **réseau profond orienté région**<sup>1</sup> pour compter des fruits *et évaluer leur maturité*

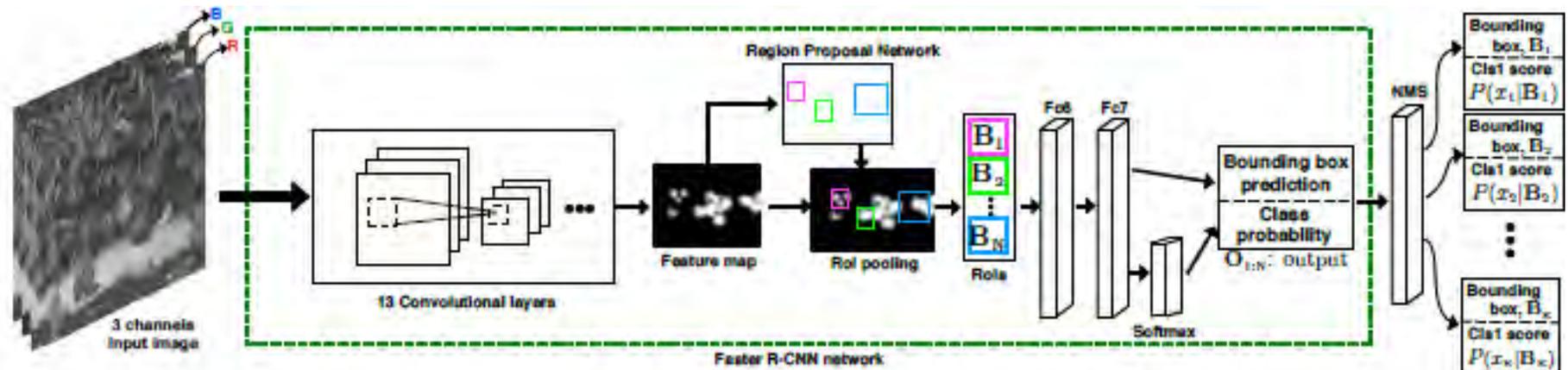


Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., & McCool, C. (2016). Deepfruits: A fruit detection system using deep neural networks. *Sensors*, 16(8), 1222.

→ **superpositions / occlusions**

→ **généricité**

## Faster Region-based Convolutional Neural Network



**Figure 3.** Illustration of test time the Faster Region-based Convolutional Neural Network (R-CNN). There are 13 convolutional and 2 fully-connected (Fc6 and Fc7) and one softmax classifier layers.  $N$  denotes the number of proposals and is set as 300.  $O_{1:N}$  is the output that contains  $N$  bounding boxes and their scores. Non-Maximum Suppression (NMS) with a threshold of 0.3 removes duplicate predictions.  $B_k$  is a bounding box of the  $K$ -th detection that is a  $4 \times 1$  vector containing the coordinates of top-left and bottom right points.  $x_k$  is a scalar representing an object being detected.

## Des données hétérogènes



Bucodiekhale (BDH) – 5 m



Keit – 10 m



Kent – 5 m

→ **variétés, conditions | distances d'acquisition**

## Des données hétérogènes



Kent – 5 m

→ stades de développement

→ positions dans l'arbre

→ illumination

### La vérité « terrain »

- sensible : concerne l'*expert*
- centrale : conditionne l'*apprentissage* du réseau
- majeure : impacte la *performance (de prédiction)* du réseau



# La validation de la « méthode » : F-mesure

→ mesurer l'écart entre les annotations expert et réseau

Annotation Expert



Prédiction Réseau

pas mangue

mangue



*Kent (Hortsys)*

**vrai positif (VP)** - **faux négatif (FN)** -  
**faux positif (FP)**

## La vérité « terrain » & experts

→ pas aussi simple à *réaliser* et à *formaliser* !



∅ 150 images | 2 annotateurs → 10% de différences (75 / 750)

### La vérité « terrain » : comment est-elle faite ?

→ un compromis entre annotations expertes et prédictions réseau

1. annotations expertes à partir de **384 images** :

→ **845 annotations validées**

2. Le réseau est *finetuné* avec **845 annotations**, *les 384 images testées*

→ *5 FN sont enlevés, 22 FP sont rajoutées*

→ **862 annotations**

3. Le réseau est *finetuné* avec **862 annotations**

## Les images couleurs



$$F\text{-mesure} = 2 \frac{\text{Rappel} \times \text{Précision}}{\text{Rappel} + \text{Précision}}$$

$$\left\{ \begin{array}{l} \text{Rappel} = \frac{|VP|}{|VP| + |FN|} \\ \text{Précision} = \frac{|VP|}{|VP| + |FP|} \end{array} \right.$$

→ F-mesure :  $\approx 0,76$

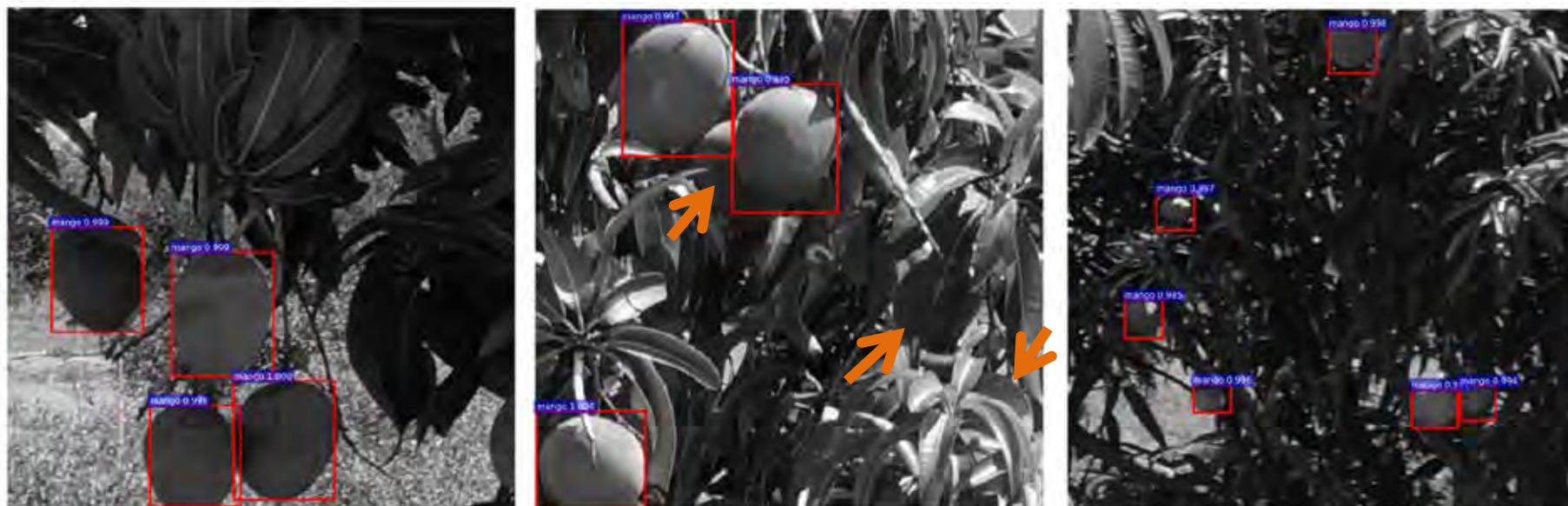
→ des proba de prédiction élevées

∅ validation croisée → 1582 annotations : Apprentissage (862) | Prédictions (730)

# Résultats

## Les images en niveau de gris

→ sensibilité du réseau à la luminance | l'intensité | la saturation



→ F-mesure :  $\approx 0,74$

∅ validation croisée → 1582 annotations : Apprentissage (862) | Prédictions (730)

# Le réseau Faster R-CNN en quelques chiffres

∅ **une carte quadro M2000**

→ *780 cœurs* | 4 GB | **650 €**

∅ **annotations expertes**

→ **800** mangues en **10 heures**

∅ **entraînement réseau**

→ **800** mangues en **6 heures**

∅ **prédiction réseau**

→ *300 ms* / imagerie 500x500

→ **28,8 s** / **image native 4000x6000**

# A propos du réseau Faster R-CNN

## ∅ apprentissage & précision de prédiction

→ *comment réaliser le jeu d'apprentissage ?*

## ∅ architecture & spécialisation du réseau

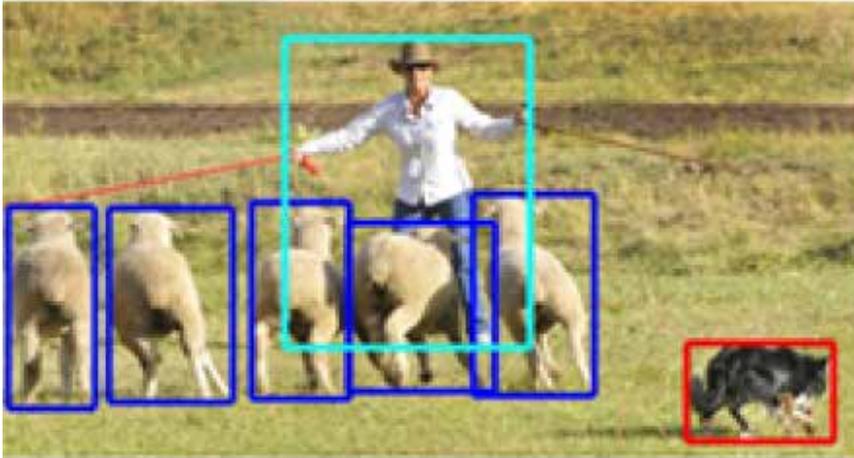
→ *comment adapter le réseau ?*

# A propos des *Mangues*

## ∅ surfaces visibles et maturité des fruits

→ *segmenter le contenu des boites*

→ *analyser la distribution spectrale de la mangue*



*Piotr Dollar, 2016, Segmenting and refining images with SharpMask, FaceBook AI research*

→ *Faster R-CNN + Fully CNN actuellement dans l'impasse*

→ *Point important : temps d'annotation ≈ 100 heures*

## Conclusion

### Le réseau Faster R-CNN probablement

∅ **transposable** à *de nombreux objets d'étude*

∅ **améliorable** *dans ses prédictions*

∅ **couplable** *pour aborder des questions de modélisation ou de diag.*

**Merci de votre attention**