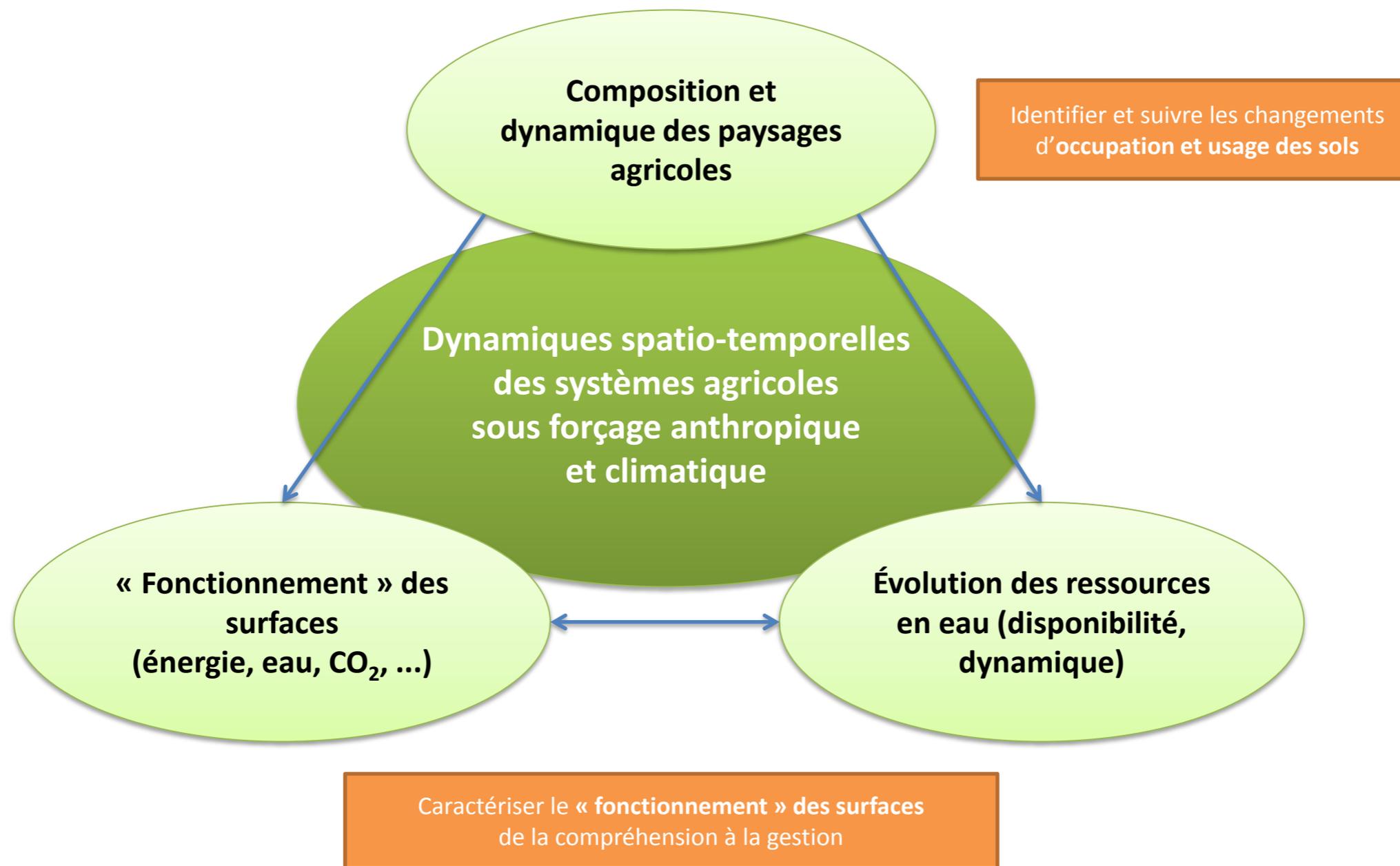


# Applications du Deep Learning à la Télédétection

**Raffaele Gaetano, Dino Ienco**  
Chargées de recherche,  
CIRAD/IRSTEA – UMR TETIS, Montpellier,  
[raffaele.gaetano@cirad.fr](mailto:raffaele.gaetano@cirad.fr)  
[dino.ienco@irstea.fr](mailto:dino.ienco@irstea.fr)

## Contexte Scientifique

- **Caractérisation et suivi des systèmes agricoles en milieu tropical**



- ❖ Pour comprendre la **durabilité de l'agriculture** sur un territoire
  - ❖ ...évaluer en temps utile la productivité d'un système agricole (agro-industrie vs. agriculture familiale, etc.)
  - ❖ ...évaluer les risques environnementaux (surexploitation des ressources en eau, érosion et dégradation des terres, etc.)
- ❖ Pour aider le **suivi agricole**
  - ❖ Sécurité alimentaire et systèmes d'alerte précoce
  - ❖ Echantillonnage de terrain, changement d'échelle
- ❖ **Système agricole ~ Paysage agricole**
  - ❖ **Télédétection comme moyen d'observation**
  - ❖ Le défi de la **diversité**
  - ❖ **Cartographie de l'occupation des sols**
  - ❖ Analyse des tendances de NDVI
  - ❖ Couplage modélisation / télédétection
  - ❖ ...



- ❖ Des cartes d'**occupation des sols pour les paysages agricoles**
  - ❖ *Besoin de précision, en particulier (mais pas que) sur le **domaine cultivé***
  - ❖ *Une **nomenclature « difficile »**, ex. des types de culture aux **pratiques agricoles***
  
- ❖ Des **données d'observation de la terre aux outils opérationnels**
  - ❖ *Données de **télé-détection** de plus en plus **variées, volumineuses et accessibles***
  - ❖ *L'**automatisation des traitements (machine learning « classique »)** comme clé pour la **production en temps utile** d'information spatiale utilisable*
    - ❖ *De l'expérience avec **iota<sup>2</sup>** [1] à la **chaîne de traitement Moringa** [2]*
  
- ❖ Des **marges d'amélioration** significatifs → un regard vers le **Deep Learning**
  - ❖ *L'occasion d'une **reflexion commune** au sein de l'UMR et du réseau de partenaires*
  - ❖ *Des études de **perspective**, sans oublier notre « vocation opérationnelle »*

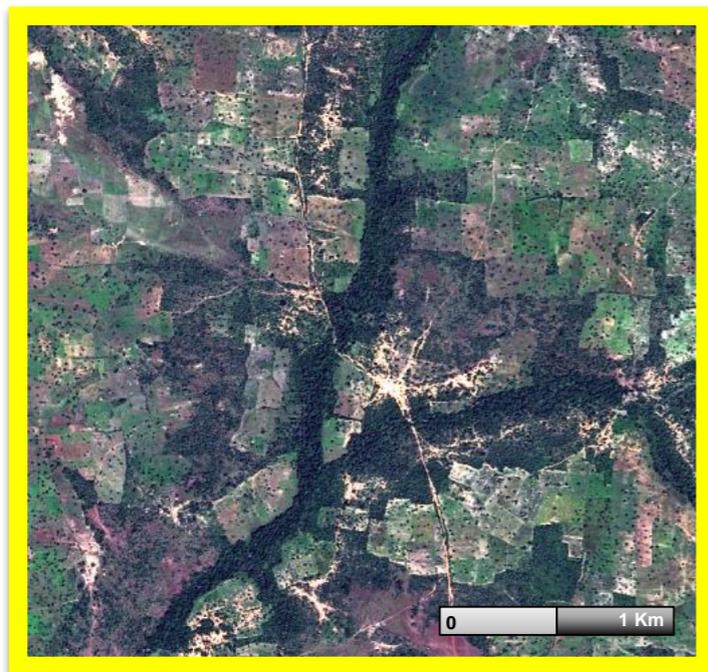
[1] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes, "Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series", *Remote Sensing*, 2017, 9(1), 95

[2] R. Gaetano et al., "The MORINGA Processing Chain: Automatic Object-based Land Cover Classification of Tropical Agrosystems using Multi-Sensor Satellite Imagery", in *ESA Living Planet Symposium (LPS 2019)*, 2019.

# Télédétection et Deep Learning

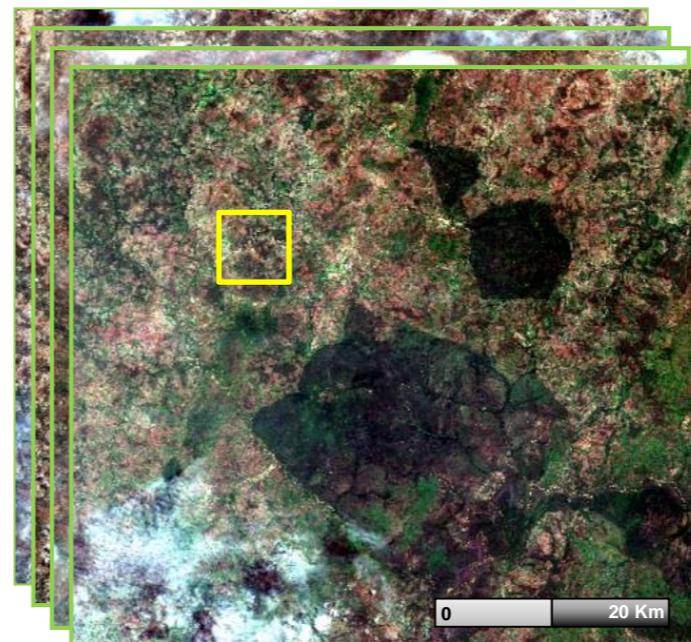
- ❖ Un **défi majeur** : exploiter l'**hétérogénéité** des données de **télédétection**
  - ❖ **Modes d'acquisition** (imagerie optique, radar, MNTs, LiDAR, etc..), **résolution spatiales** (des 100m au sub-métrique), **fréquences d'acquisition** (de qq jours à un an ou plus)

- ❖ Un **défi majeur** : exploiter l'**hétérogénéité** des données de **télédétection**
  - ❖ **Modes d'acquisition** (imagerie optique, radar, MNTs, LiDAR, etc.), **résolution spatiales** (des 100m au sub-métrique), **fréquences d'acquisition** (de qq jours à un an ou plus)
- ❖ Un « **dualisme** » particulièrement intéressant : **THRS** vs. **STIS**



**Très Haute Résolution Spatiale**  
pixels de ~1m (ex. SPOT6/7),  
une ou deux acquisitions environ par an

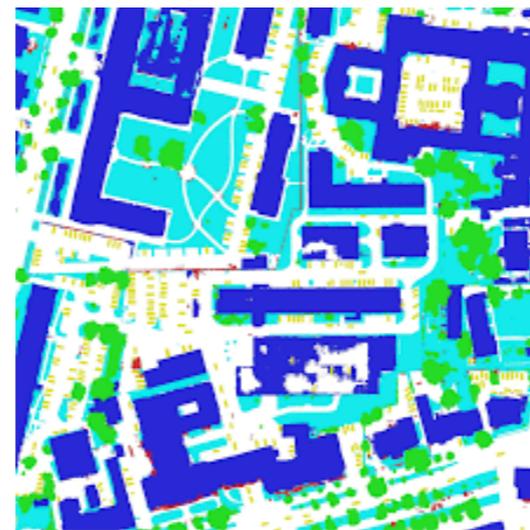
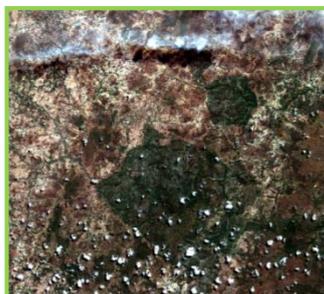
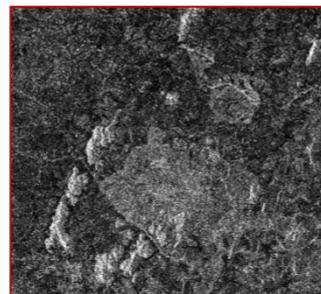
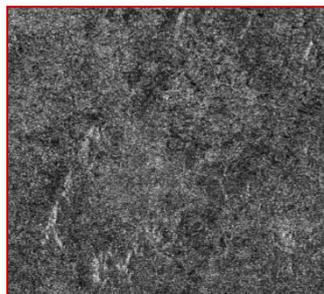
Utile pour caractériser les structures d'un paysage à l'échelle fine et les « motifs » spatiaux



**Séries Temporelles d'Images Satellites**  
pixels de ~10m (ex. Sentinel-2/-1),  
une acquisition tous les 5 (6) jours

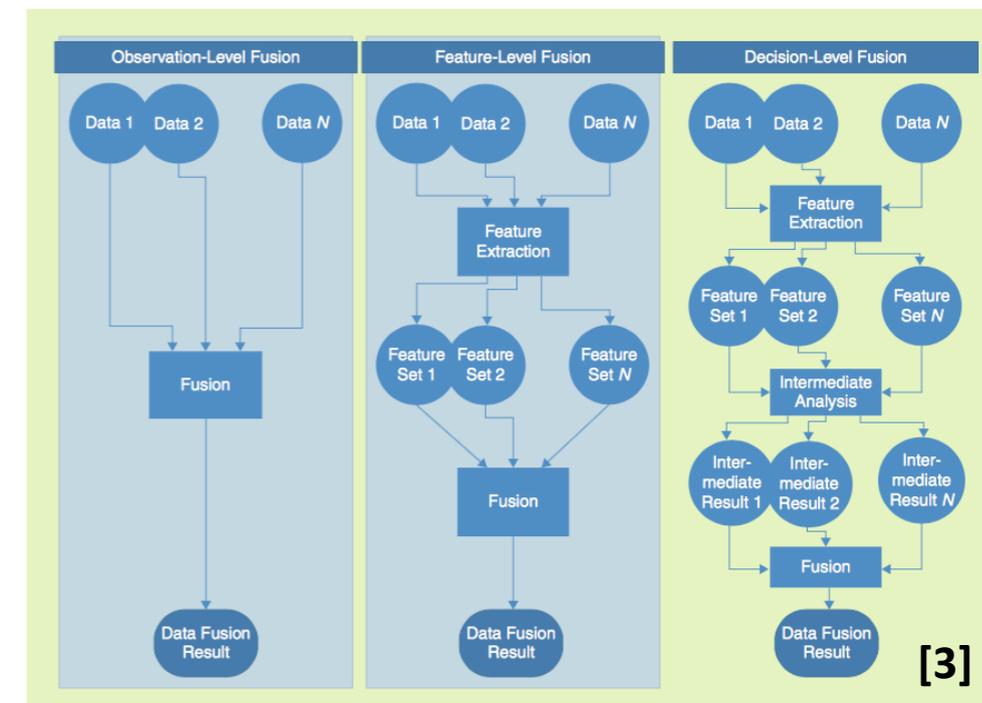
Utile pour le suivi des dynamiques spatio-temporelles d'un paysage (tendances et anomalies)

- ❖ Un **défi majeur** : exploiter l'**hétérogénéité** des données de **télédétection**
  - ❖ *Modes d'acquisition (imagerie optique, radar, MNTs, LiDAR, etc.), résolutions spatiales (des 100m au sub-métrique), fréquences d'acquisition (de qq jours à un an ou plus)*
- ❖ Un « **dualisme** » particulièrement intéressant : **THRS** vs. **STIS**
- ❖ Des **contraintes** vis-à-vis de l'opérationnalisation
  - ❖ *Qualité des images (nébulosité, niveaux de traitement, etc.)*
  - ❖ *Quantité et qualité de la donnée de référence*



## ❖ Hétérogénéité → approches par fusion de données

- ❖ Pour l'**occupation des sols**, en *machine learning* classique les stratégies de fusion **aux niveaux des caractéristiques** (*feature-level data fusion*) sont privilégiées...
- ❖ ...mais le **volume** et la **diversité** rendent les processus d'extraction de caractéristiques difficiles à optimiser



## ❖ En Deep Learning, les processus d'extraction des *features* sont intégrés dans les mécanismes d'apprentissage (*representational learning*)

- ❖ *L'extraction d'informations des données est automatique et optimisée en fonction de la tâche*

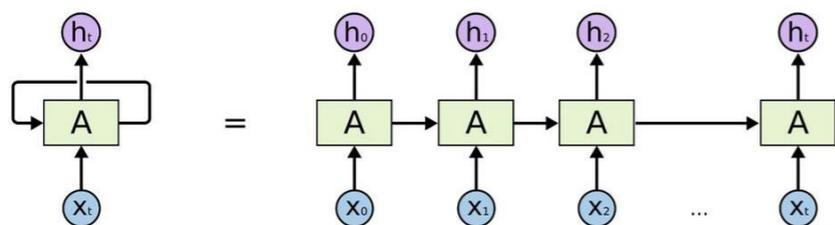
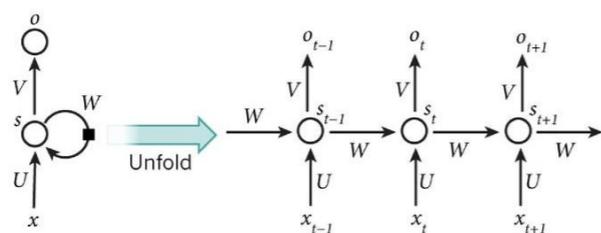
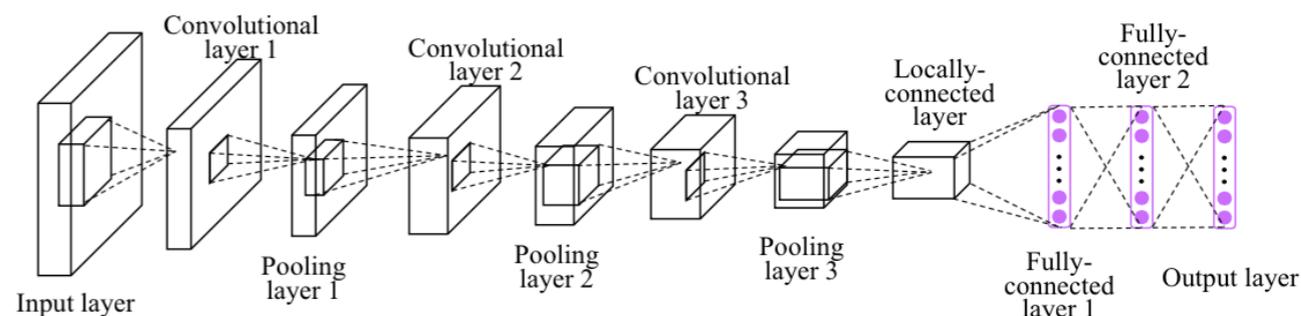


[3] M. Schmitt and X. X. Zhu, "Data Fusion and Remote Sensing: An ever-growing relationship". IEEE Geoscience and Remote Sensing Magazine 4(4): 6--23, 2016

- ❖ **Dualité THRS / STIS → Complémentarité des dimensions spatiale et temporelle**
- ❖ L'état de l'art en **Deep Learning** met à disposition des « **briques** » d'apprentissage adaptés à l'analyse de ces dimensions dont l'**usage combiné** est facile de mise en œuvre

## CNNs (Convolutional Neural Networks)

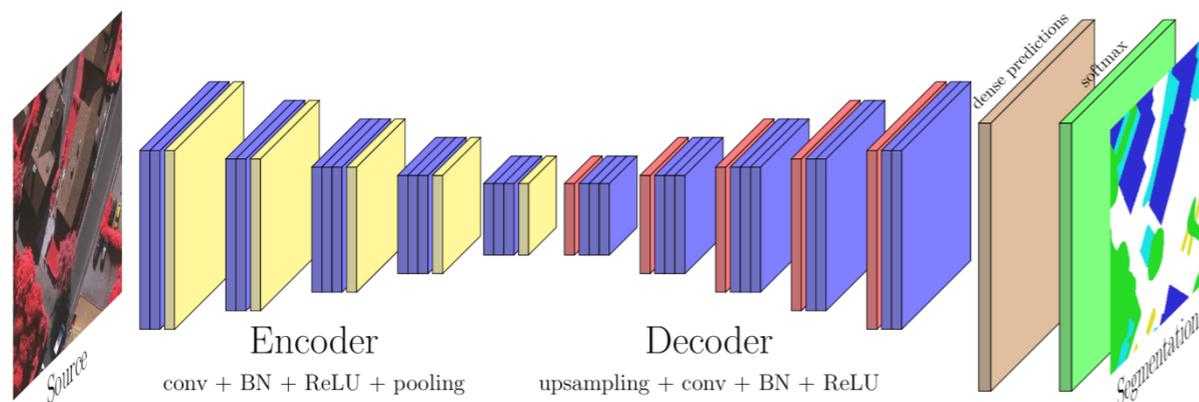
Inspirés de la vision humaine, les **réseaux de neurones convolutifs** permettent la prise en compte des (auto) corrélations spatiales (contexte, voisinages structures et textures)



## RNNs (Recurrent Neural Networks)

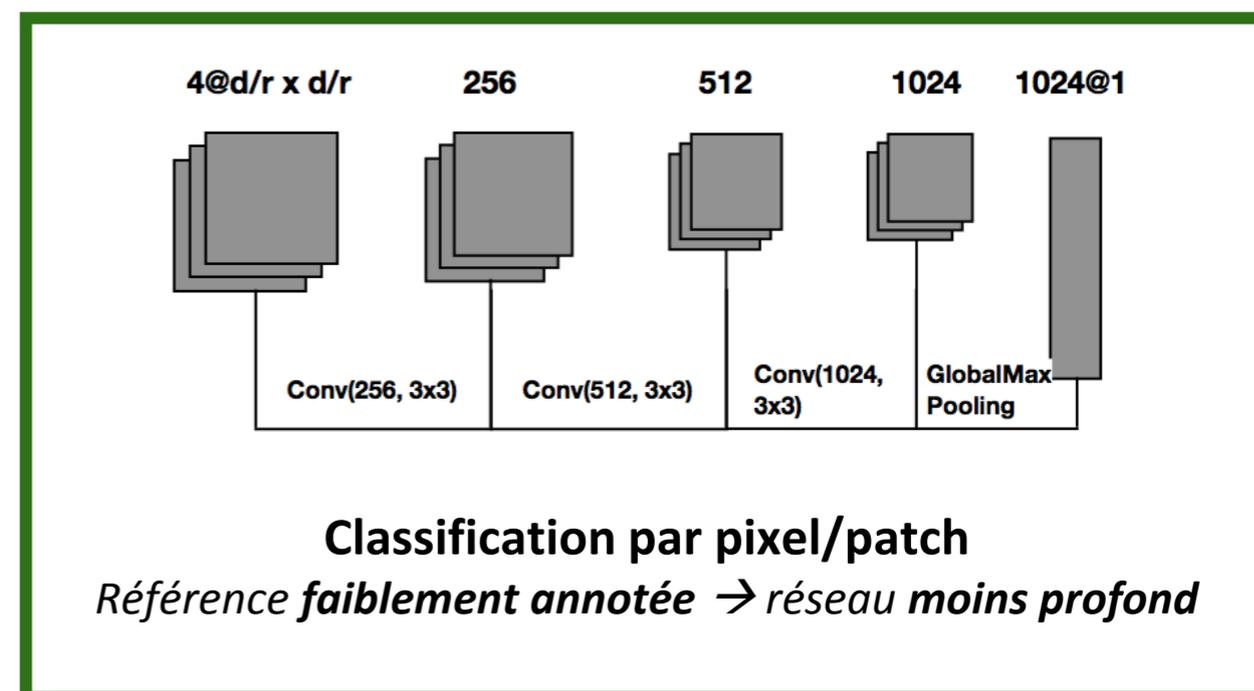
Très utilisés dans l'analyse du langage, les **réseaux de neurones récurrents** sont conçus pour tirer partie de la corrélation dans des suites de valeurs (tendances, anomalies, interdépendances temporelles)

- ❖ **Quantité de données de référence (références « faiblement » annotées)**
- ❖ La « gourmandise » du **Deep Learning** en termes de **données d'apprentissage** est une idée reçue répandue, en partie vraie, mais cela dépend de plusieurs facteurs:
  - ❖ **complexité de la tâche** (ex. nombre de classes et typologie pour l'OcSol)
  - ❖ **pertinence et qualité des données en entrée**
  - ❖ **complexité de l'architecture (ex. profondeur)**
- ❖ ...la **littérature émergente en TLD** promut, entre autres, les approches dites *shallow*



## Segmentation sémantique

Référence **densément annotée** → réseau **plus profond**



## Classification par pixel/patch

Référence **faiblement annotée** → réseau **moins profond**

## Exemples d'applications :

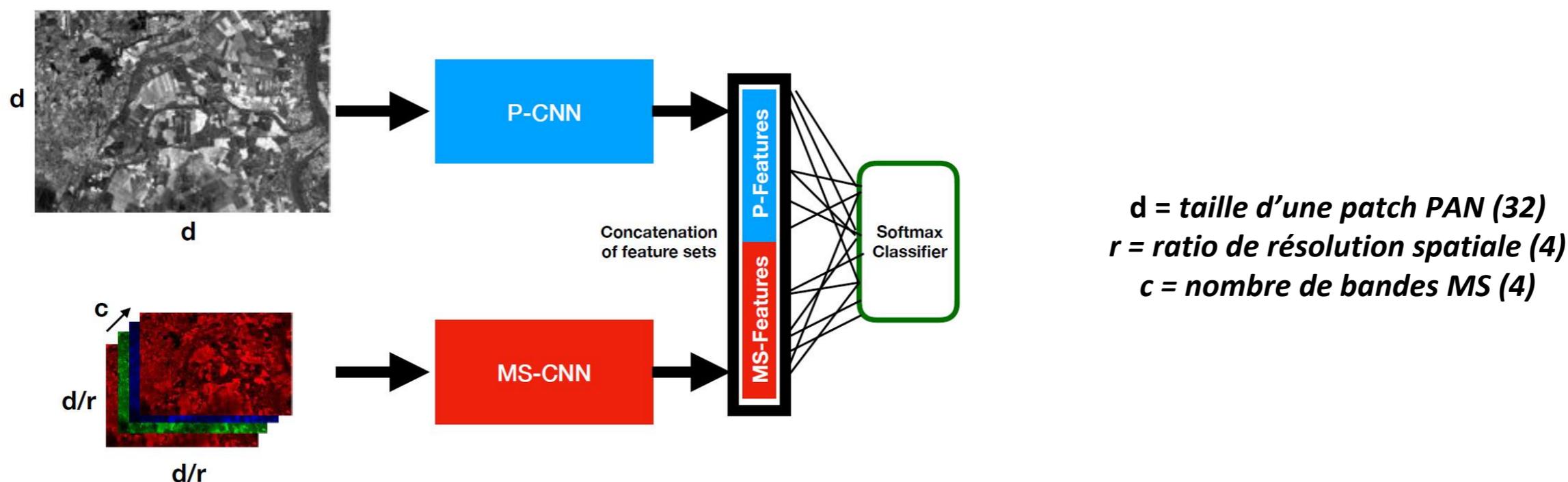
# Méthodologies basées sur Deep Learning pour la classification de l'Occupation du Sol

## ***MultiResoLCC*** : une approche **Deep Learning** pour la **classification de l'occupation du sol** à partir d'imagerie **multi-résolution** [4]

- ❖ Une scène à **THRS** (ex. SPOT6/7, Pléiades) se compose typiquement de deux images
  - ❖ Une **bande panchromatique (PAN)** à résolution spatiale fine (1.5m pour SPOT6/7)
  - ❖ Une **image multispectrale (MS)** à résolution moins fine (6m pour SPOT6/7)
- ❖ **Objectifs** : concevoir une solution DL pour la classification de l'OcSol à partir d'une image à **très haute résolution spatiale** dans son **format natif**
  - ❖ Un scénario de **fusion mono-capteur** [3]
  - ❖ Exploiter les **corrélations spatiales à deux échelles** et éviter le *pan-sharpening*
- ❖ **Méthodologie** : définition d'une architecture à **deux branches opérant en parallèle sur les deux sources disponibles** (PAN et MS), basée sur des briques de type **convolutif**

[4] R. Gaetano, D. Ienco, K. Ose, R. Cresson: "A Two-Branch CNN Architecture for Land Cover Classification of PAN and MS Imagery". Remote Sensing 10(11): 1746 (2018)

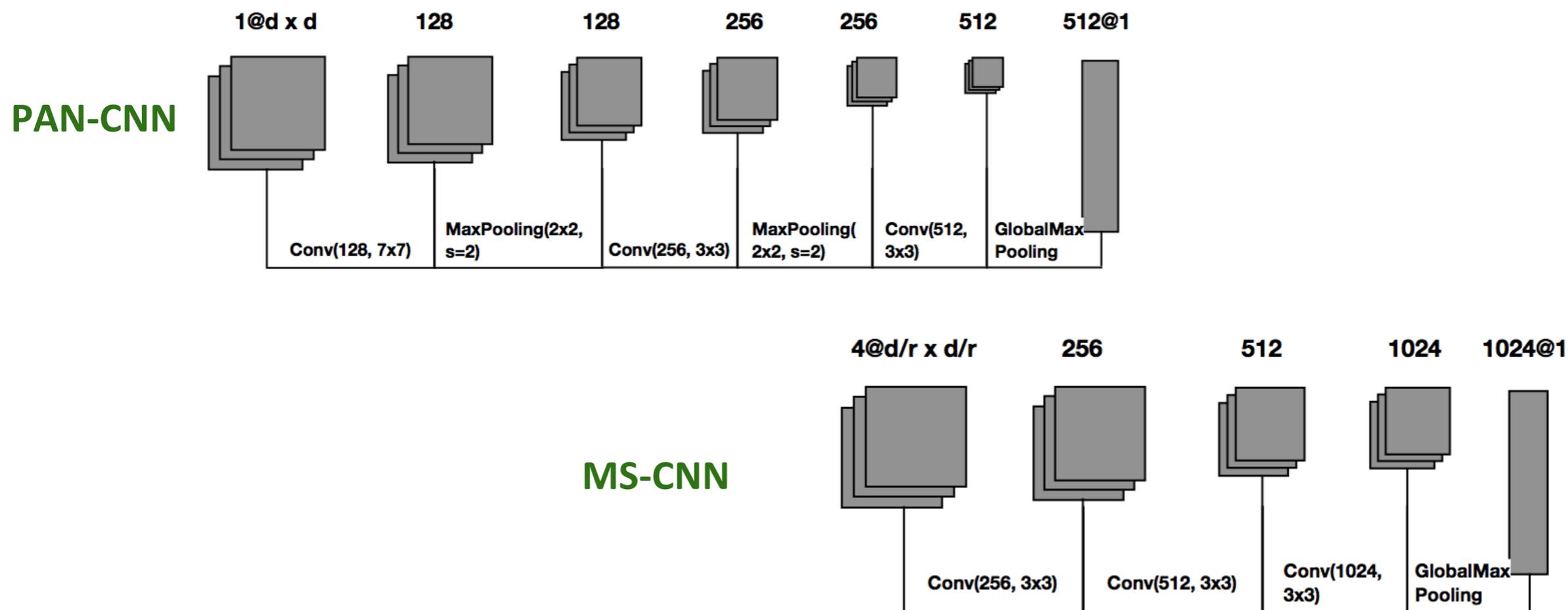
## *MultiResoLCC* : une approche Deep Learning pour la classification de l'occupation du sol à partir d'imagerie multi-résolution [4]



- ❖ Architecture « *end-to-end* », apprentissage « *from scratch* »
- ❖ Un **module CNN dédié** pour chaque source, avec **concaténation des features** extraites (*late fusion*)
- ❖ La **classification** est produite à la **résolution de la PAN** (1.5m)

[4] R. Gaetano, D. Ienco, K. Ose, R. Cresson: "A Two-Branch CNN Architecture for Land Cover Classification of PAN and MS Imagery". Remote Sensing 10(11): 1746 (2018)

## *MultiResoLCC* : une approche Deep Learning pour la classification de l'occupation du sol à partir d'imagerie multi-résolution [4]



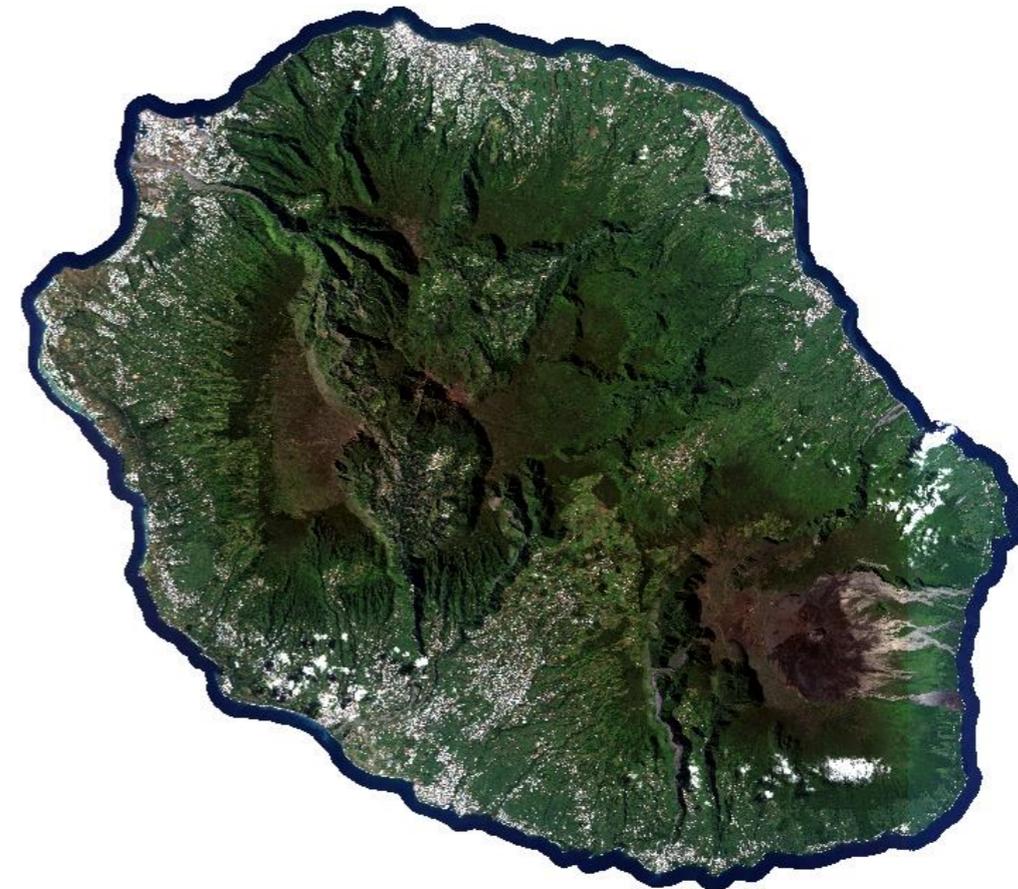
- ❖ Fonctions d'activation de type **ReLU (Rectified Linear Unit)**
- ❖ Utilisation de la **normalisation par batch**

[4] R. Gaetano, D. Ienco, K. Ose, R. Cresson: "A Two-Branch CNN Architecture for Land Cover Classification of PAN and MS Imagery". Remote Sensing 10(11): 1746 (2018)

## *MultiResoLCC* : jeux de données



**Un site en France métropolitaine**  
*Zone Sud-Gard (est de Montpellier)*



**Un site en milieu tropical**  
*Île de la Réunion*

## MultiResoLCC : jeux de données

- Une scene **SPOT6/7** (Mars 2016)
- Taille **PAN** : **24110 x 33740** pixels
- Taille **MS** : **6028 x 8435** pixels
- Une nomenclature à **8 classes**

Class	Label	# Objects	# Pixels
1	<i>Cereal Crops</i>	167	50100
2	<i>Other Crops</i>	167	50098
3	<i>Tree Crops</i>	167	50027
4	<i>Meadows</i>	167	49997
5	<i>Vineyard</i>	167	50100
6	<i>Forest</i>	172	50273
7	<i>Urban areas</i>	222	50275
8	<i>Water Surfaces</i>	167	50100

**Un site en France métropolitaine**  
*Zone Sud-Gard (est de Montpellier)*

- Une scene **SPOT6/7** (Avril 2016)
- Taille **PAN** : **44374 x 39422** pixels
- Taille **MS** : **11094 x 9856** pixels
- Une nomenclature à **13 classes**

Class	Label	# Objects	# Pixels
1	<i>Crop Cultivations</i>	168	50061
2	<i>Sugar cane</i>	167	50100
3	<i>Orchards</i>	167	50092
4	<i>Forest plantations</i>	67	20100
5	<i>Meadow</i>	167	50100
6	<i>Forest</i>	167	50100
7	<i>Shrubby savannah</i>	173	50263
8	<i>Herbaceous savannah</i>	78	23302
9	<i>Bare rocks</i>	107	31587
10	<i>Urban areas</i>	125	36046
11	<i>Greenhouse crops</i>	49	14387
12	<i>Water Surfaces</i>	96	2711
13	<i>Shadows</i>	38	11400

**Un site en milieu tropical**  
*Île de la Réunion*

## **MultiResoLCC** : cadre expérimental

- ❖ *Découpage de la donnée de référence :*
  - ❖ **30% des polygones utilisés pour l'apprentissage, 70% pour le test**
  - ❖ **Résultats moyennés sur 10 découpages aléatoires**
  
- ❖ *Comparaisons réalisées :*
  - ❖ **Random Forest appliqué aux pixels et aux patch**
  - ❖ **Une seule branche CNN appliqué à l'image après pan-sharpening**
  - ❖ **DMIL [5], une méthode DL alternative pour la même tâche**
  
- ❖ *Indicateurs de précisions utilisées :*
  - ❖ **Précision globale**
  - ❖ **F-Measure (globale et par classe)**
  - ❖ **Kappa de Cohen**

[5] X. Liu *et al.* : Deep Multiple Instance Learning-Based Spatial-Spectral Classification for PAN and MS Imagery. IEEE Trans. Geoscience and Remote Sensing 56(1): 461-473 (2018)

## MultiResoLCC : résultats

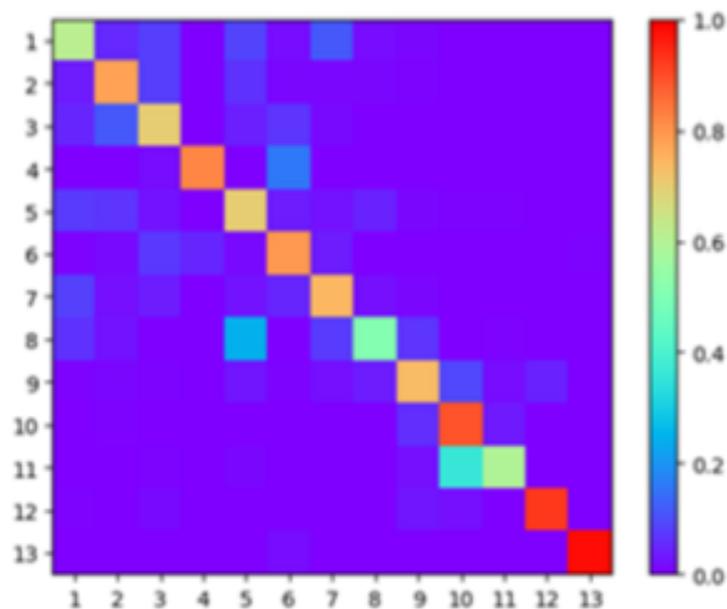
	<i>Accuracy</i>	<i>F-Measure</i>	<i>Kappa</i>
RF(PIXEL)	24.87 ± 0.2	23.66 ± 0.2	0.1719 ± 0.0024
RF(PATCH)	72.22 ± 1.31	71.53 ± 1.4	0.6943 ± 0.0144
CNN <sub>PS</sub>	74.49 ± 1.20	74.25 ± 1.24	0.7195 ± 0.0131
DMIL	69.40 ± 1.11	69.34 ± 1.12	0.6637 ± 0.0121
MRFusion	<b>79.65 ± 0.87</b>	<b>79.56 ± 0.91</b>	<b>0.7764 ± 0.0096</b>

## Site Réunion

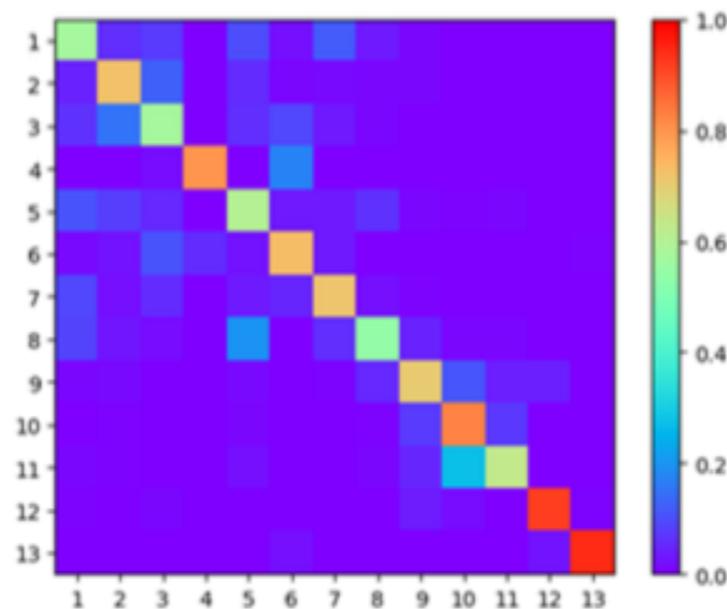
	<i>Accuracy</i>	<i>F-Measure</i>	<i>Kappa</i>
RF(PIXEL)	25.91 ± 0.16	25.52 ± 0.11	0.1532 ± 0.18
RF(PATCH)	69.93 ± 0.76	69.55 ± 0.77	0.6564 ± 0.87
CNN <sub>PS</sub>	66.14 ± 0.78	65.80 ± 0.77	0.6131 ± 0.0089
DMIL	61.96 ± 1.00	61.76 ± 1.01	0.5652 ± 0.0115
MRFusion	<b>70.48 ± 0.55</b>	<b>70.19 ± 0.67</b>	<b>0.6627 ± 0.0063</b>

## Site Gard

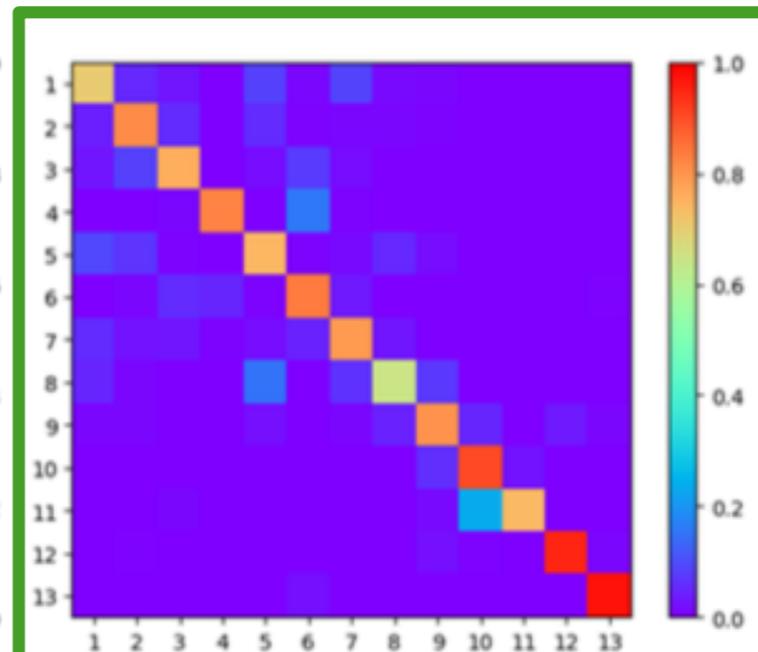
## MultiResoLCC : résultats



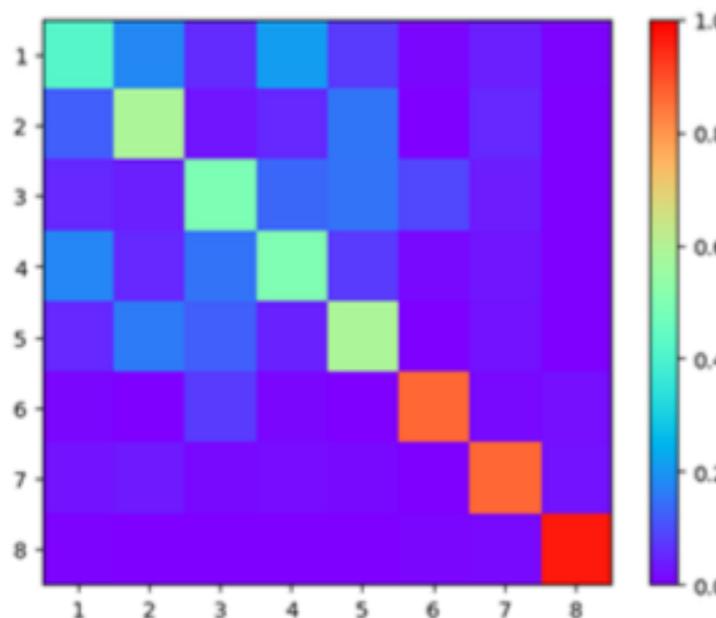
(a)



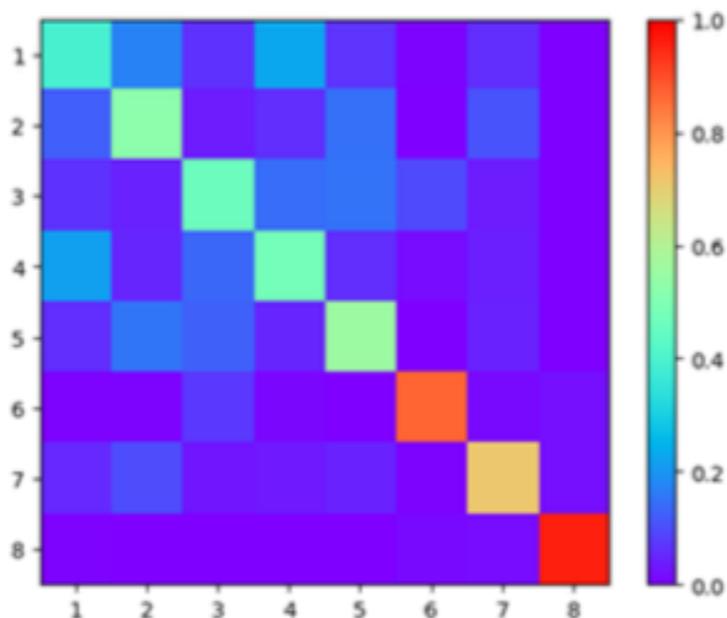
(b)



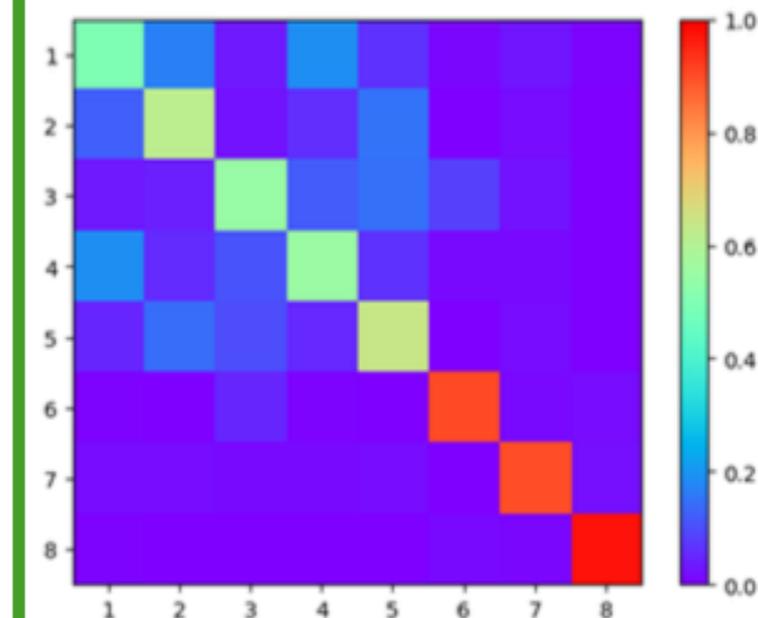
(c)



(d)

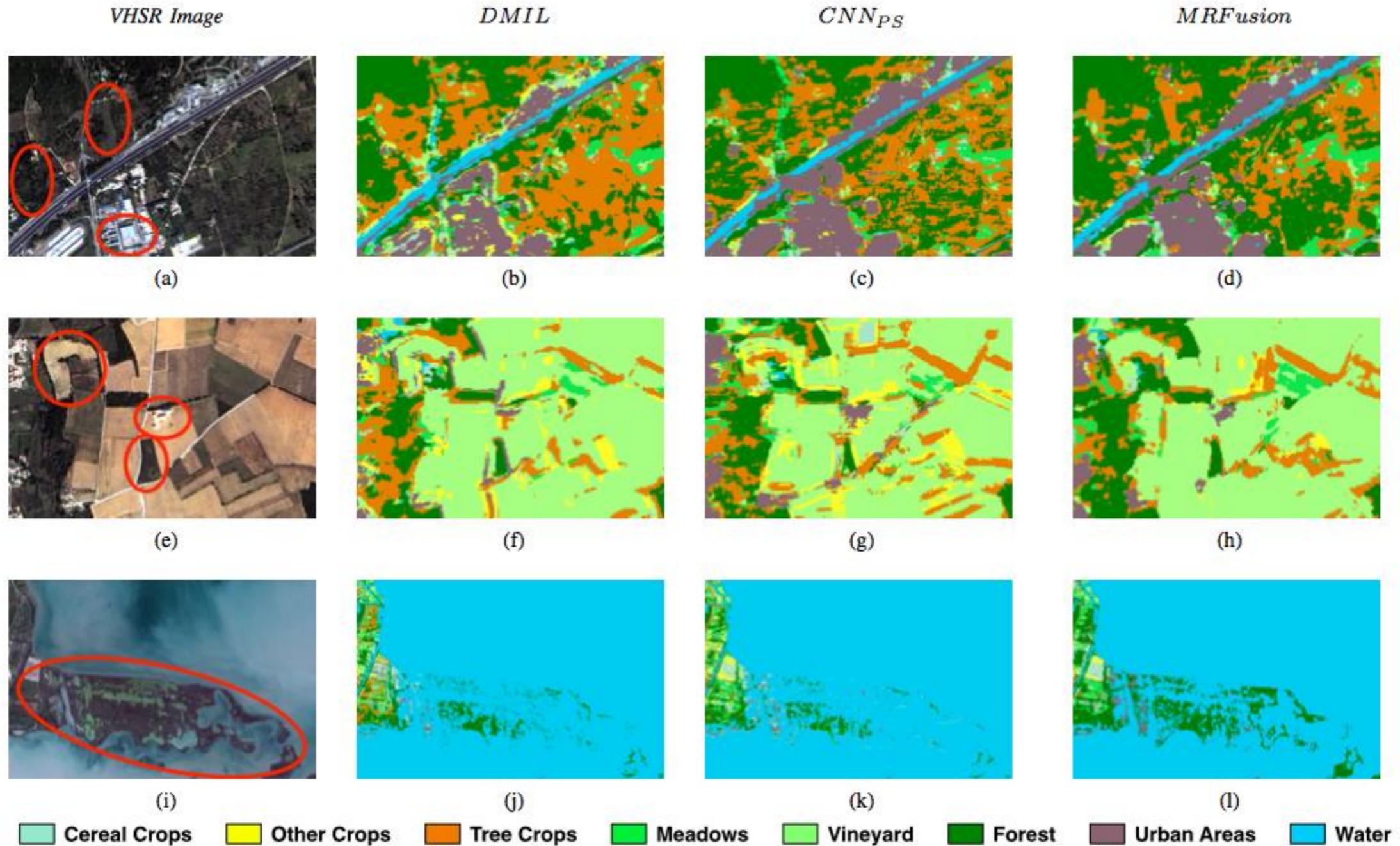


(e)

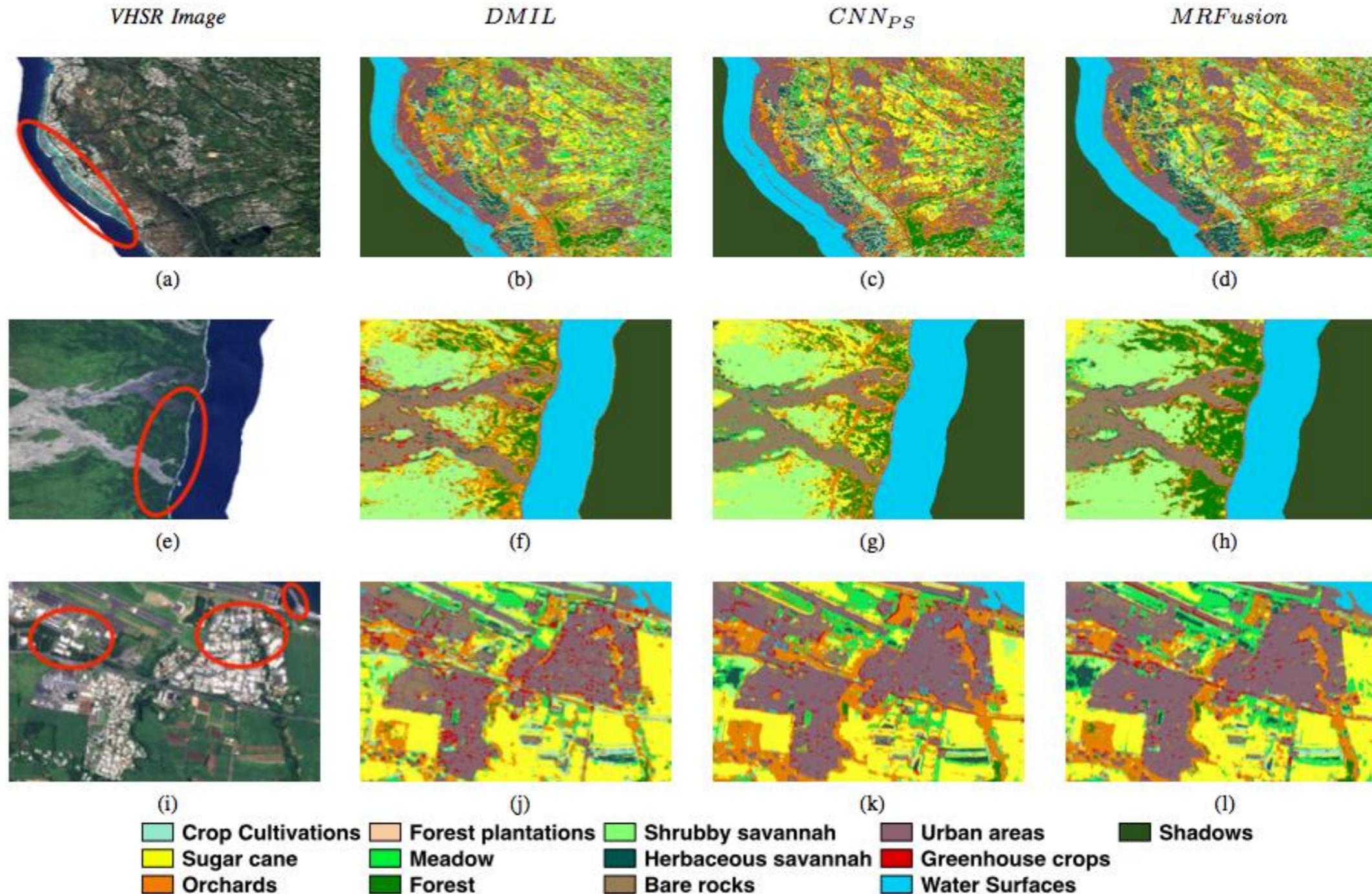


(f)

## MultiResoLCC : résultats



## MultiResoLCC : résultats



## Site Réunion

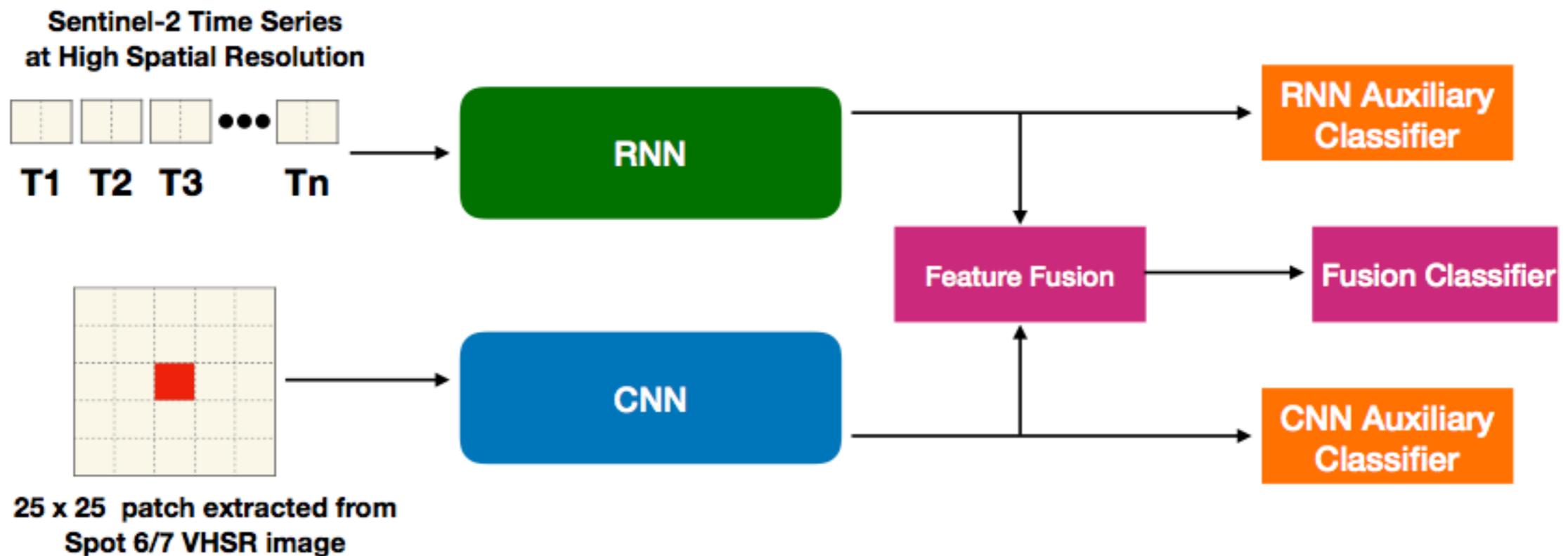
## ***M<sup>3</sup>Fusion*** : une architecture **Deep Learning** pour la **fusion d'images satellites multi-{échelles, modalités, temporelles}** [6]

- ❖ Un travail antérieur nous a permis de montrer l'apport des **RNNs en Deep Learning** pour la **classification de l'OcSol** à partir de **séries temporelles Sentinel-2** [7]
  - ❖ **Question** : quelles pistes en Deep Learning pour améliorer ce type d'approches par l'intégration d'une source à très haute résolution spatiale ?
- ❖ **Objectifs** : concevoir une solution DL pour la classification de l'OcSol à partir du couplage d'une scène à **THRS** et une **série temporelle d'images à résolution décimétrique**
  - ❖ Un premier scénario de **fusion multi-capteur** [3]
  - ❖ Analyse parallèle des **dynamiques temporelles (S2)** et des **contextes spatiaux (THRS)**
- ❖ **Méthodologie** : encore une architecture à **deux branches opérant en parallèle sur les deux sources disponibles**, basée sur des briques de type **CNN (pour THR)** et **RNN (pour S2)**

[6] P. Benedetti *et al.*, "M3Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion," IEEE J. of Sel. Top. in App. Earth Obs. and Rem. Sens., vol. 11, pp. 4939–4949, Dec 2018.

[7] D. Ienco, *et al.*, "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks," IEEE Geoscience and Remote Sensing Letters, vol. 14, pp. 1685–1689, Oct 2017.

## *M<sup>3</sup>Fusion* : une architecture Deep Learning pour la fusion d'images satellites multi-{échelles, modalités, temporelles} [6]

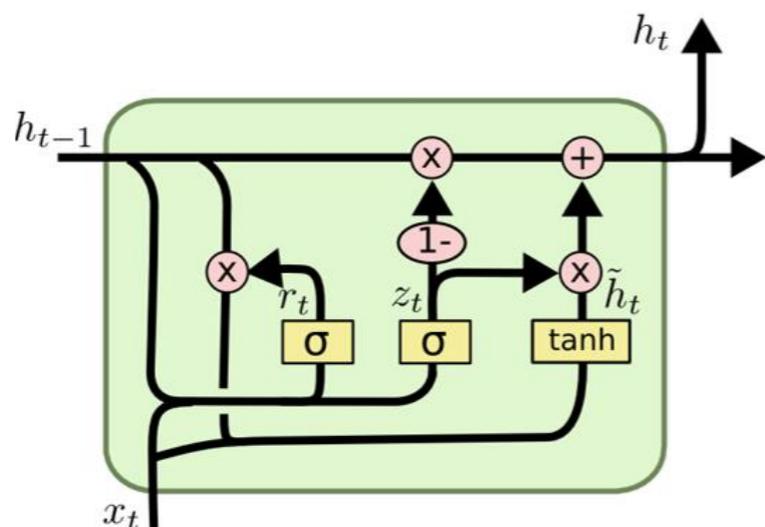


- ❖ *Un échantillon correspond à un pixel Sentinel-2 (carte à 10m) et consiste en :*
  - ❖ *une série temporelle de pixels Sentinel-2 (10 réflectances + 6 indices)*
  - ❖ *la patch de l'image THRS recouvrant une zone autour du pixel S2 correspondant*
- ❖ *Les classifieurs auxiliaires servent à forcer l'extraction de caractéristiques complémentaires [8]*

[6] P. Benedetti *et al.*, "M3Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion," IEEE J. of Sel. Top. in App. Earth Obs. and Rem. Sens., vol. 11, pp. 4939–4949, Dec 2018.

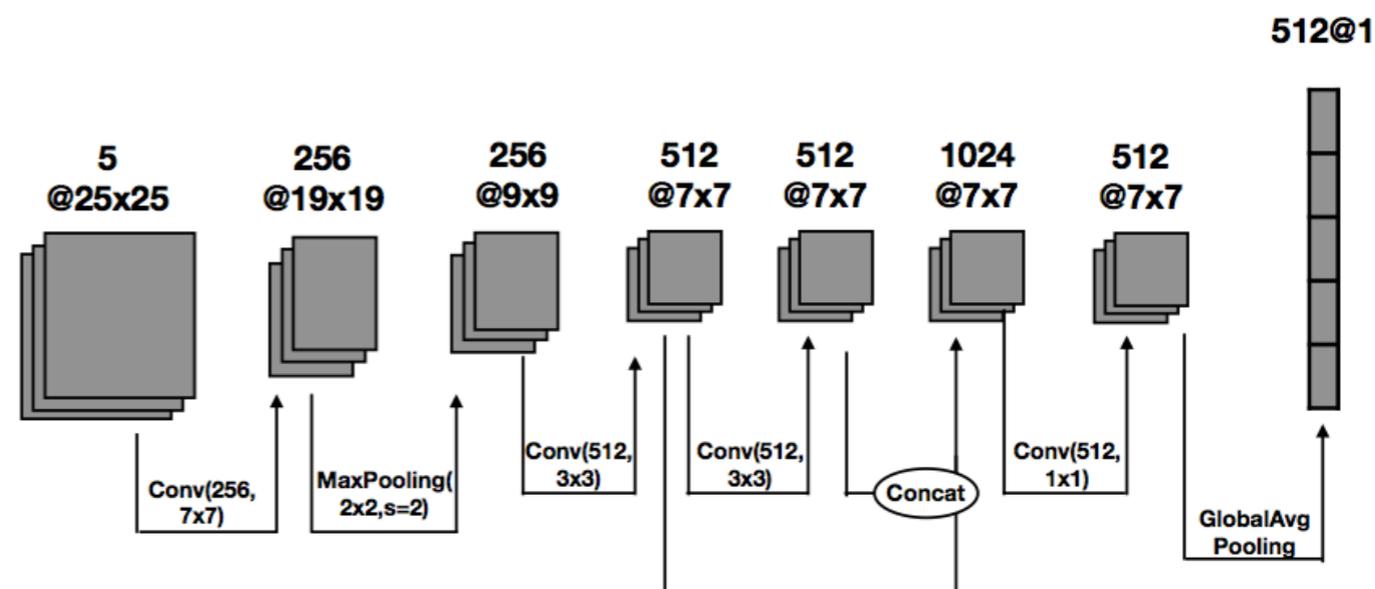
[8] S. Hou, X. Liu, and Z. Wang, "Dualnet: Learn complementary features for image recognition," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 502–510.

## *M<sup>3</sup>Fusion* : une architecture Deep Learning pour la fusion d'images satellites multi-**{échelles, modalités, temporelles}** [6]



- ❖ *Brique RNN : la Gate Recurrent Unit*
- ❖ *Plus rapide et performante comparée à d'autres (ex. LSTM)*
- ❖ *Utilisée avec un mécanisme d'attention permettant de combiner l'information extraite à chaque estampe temporelle*

- ❖ *Brique CNN:*
- ❖ *Quelque peu plus complexe que celle utilisée pour MultiResoLCC*
- ❖ *Utilisation du drop-out*



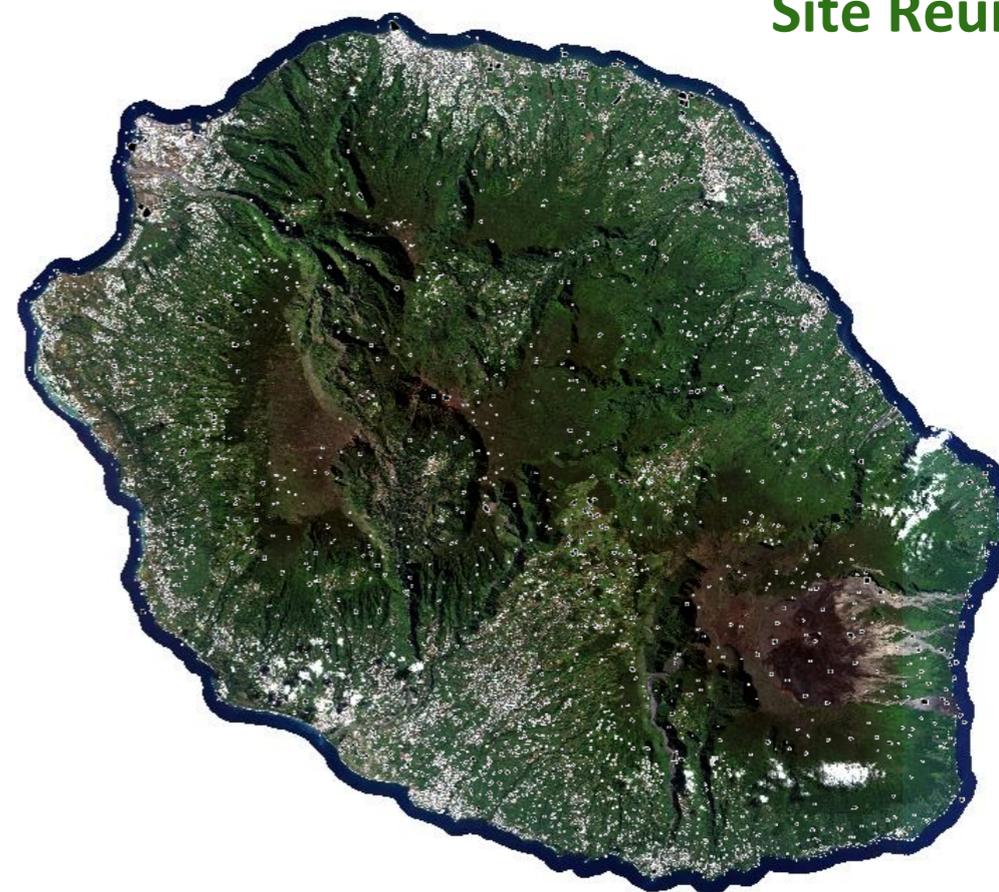
[6] P. Benedetti *et al.*, "M3Fusion: A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion," IEEE J. of Sel. Top. in App. Earth Obs. and Rem. Sens., vol. 11, pp. 4939–4949, Dec 2018.

[8] S. Hou, X. Liu, and Z. Wang, "Dualnet: Learn complementary features for image recognition," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 502–510.

## M<sup>3</sup>Fusion : jeux de données

### Site Réunion

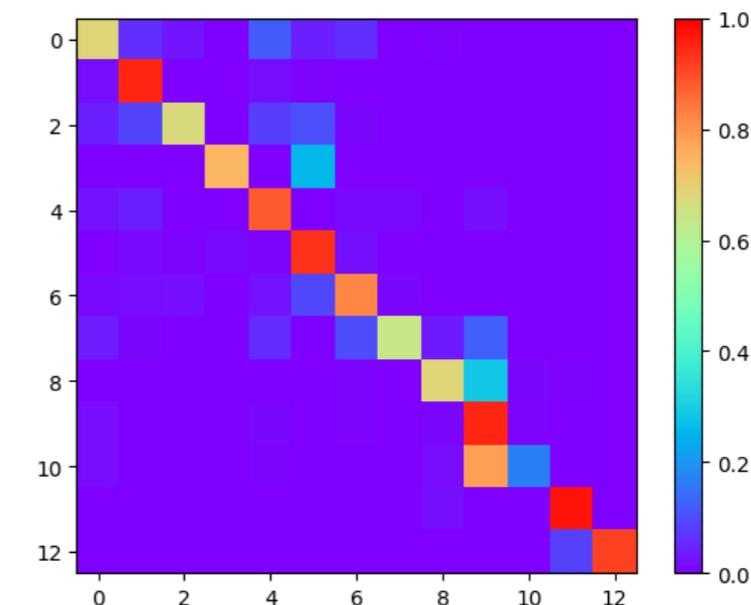
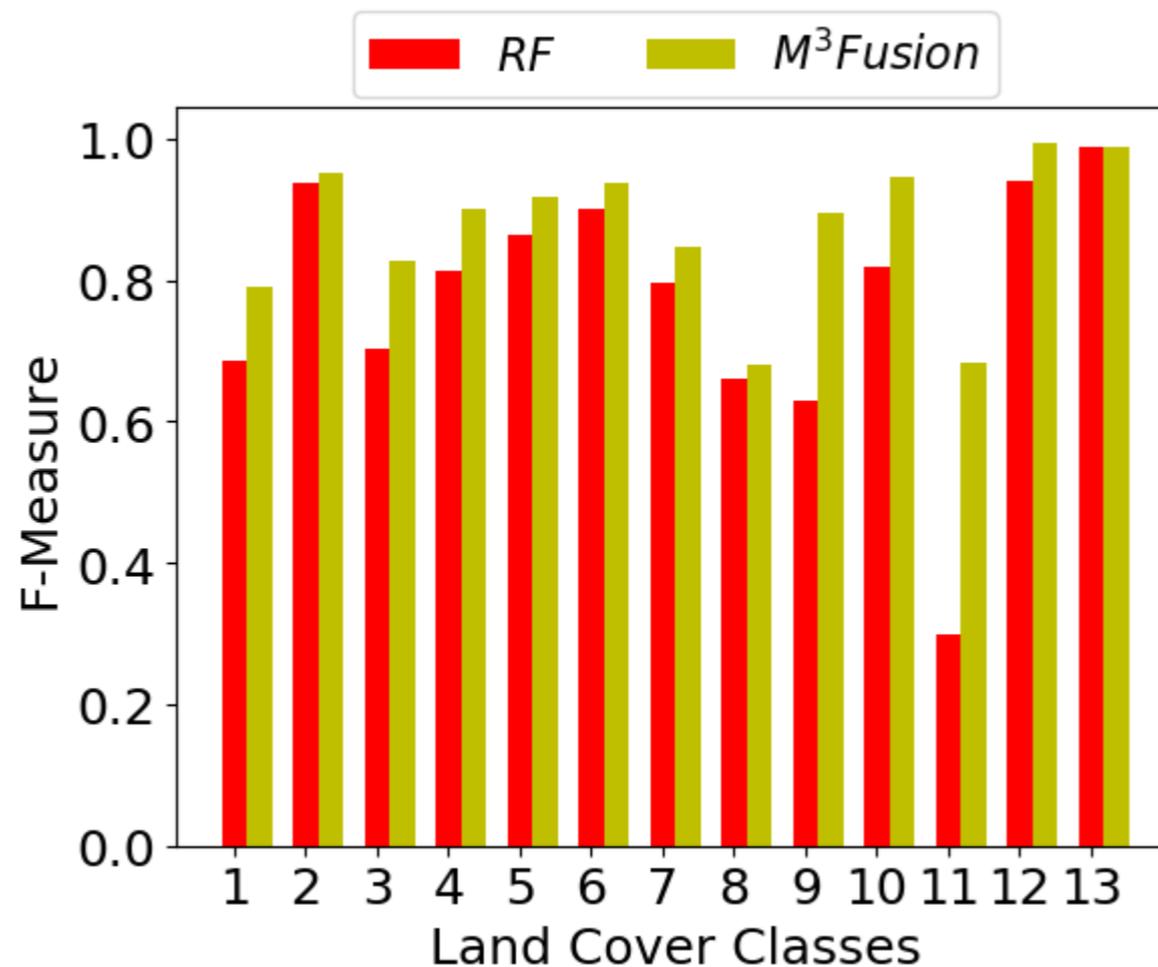
Class	Label	# Objects	# Pixels
1	<i>Crop Cultivations</i>	380	12090
2	<i>Sugar cane</i>	496	84136
3	<i>Orchards</i>	299	15477
4	<i>Forest plantations</i>	67	9783
5	<i>Meadow</i>	257	50596
6	<i>Forest</i>	292	55108
7	<i>Shrubby savannah</i>	371	20287
8	<i>Herbaceous savannah</i>	78	5978
9	<i>Bare rocks</i>	107	18659
10	<i>Urbanized areas</i>	125	36178
11	<i>Greenhouse crops</i>	50	1877
12	<i>Water Surfaces</i>	96	7349
13	<i>Shadows</i>	38	5230



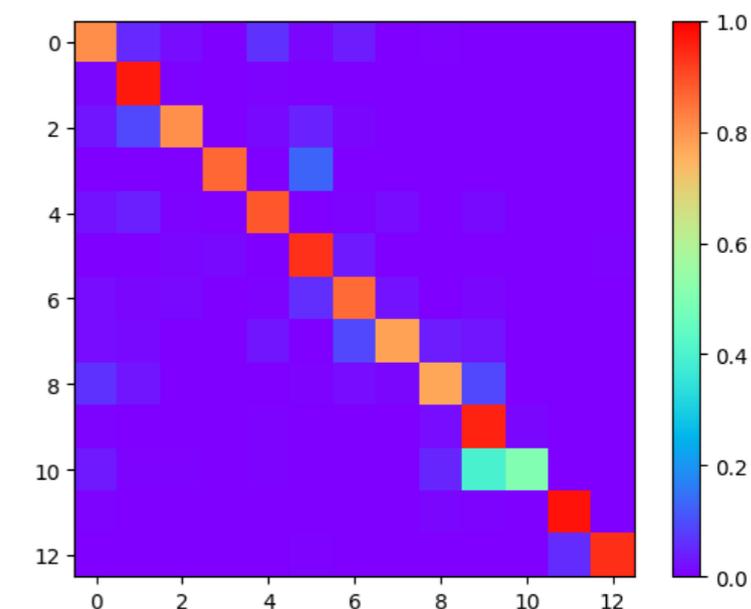
- ❖ Une série de 34 images Sentinel-2 (THEIA, Avril 2016 – Mai 2017), à 10m de résolution, chacune comptant 10 bandes et 6 indices radiométriques, « dénuagées » par **gapfilling** (interpolation temporelle)
- ❖ Une image SPOT6/7 (Avril 2016) à 1.5m de résolution (pan-sharpening)
- ❖ Référence basée sur une campagne de terrain et intégrée avec des BD existantes (RPG) et photo-interprétation

## *M<sup>3</sup>Fusion* : résultats

Class	Label
1	<i>Crop Cultivations</i>
2	<i>Sugar cane</i>
3	<i>Orchards</i>
4	<i>Forest plantations</i>
5	<i>Meadow</i>
6	<i>Forest</i>
7	<i>Shrubby savannah</i>
8	<i>Herbaceous savannah</i>
9	<i>Bare rocks</i>
10	<i>Urbanized areas</i>
11	<i>Greenhouse crops</i>
12	<i>Water Surfaces</i>
13	<i>Shadows</i>



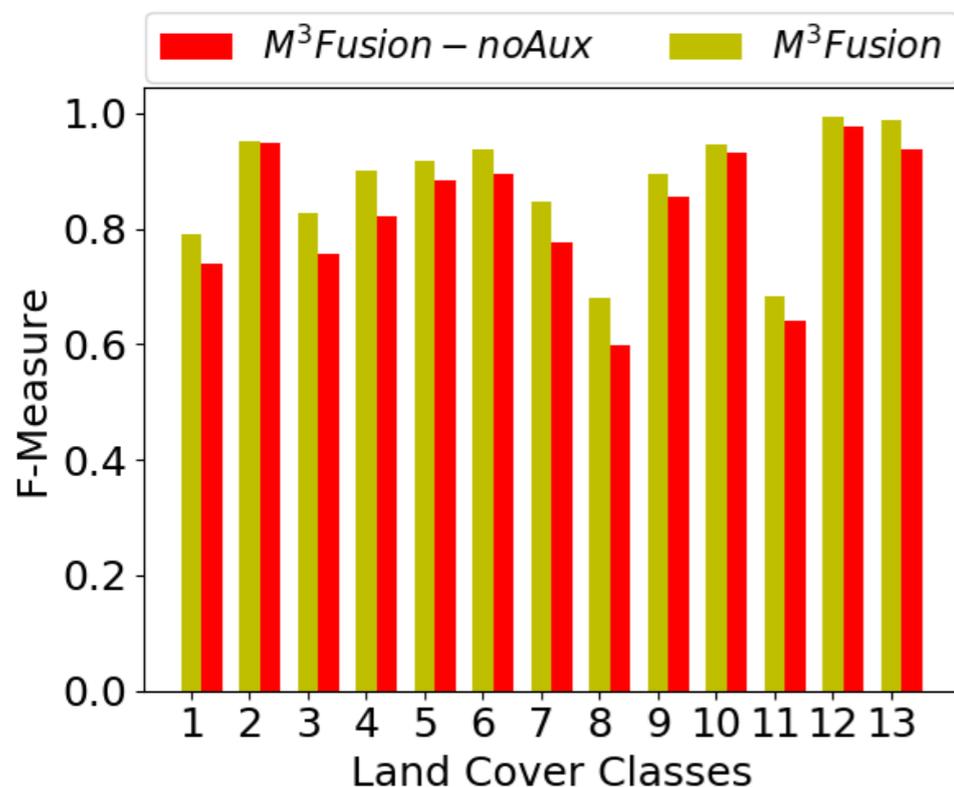
*Random Forest*



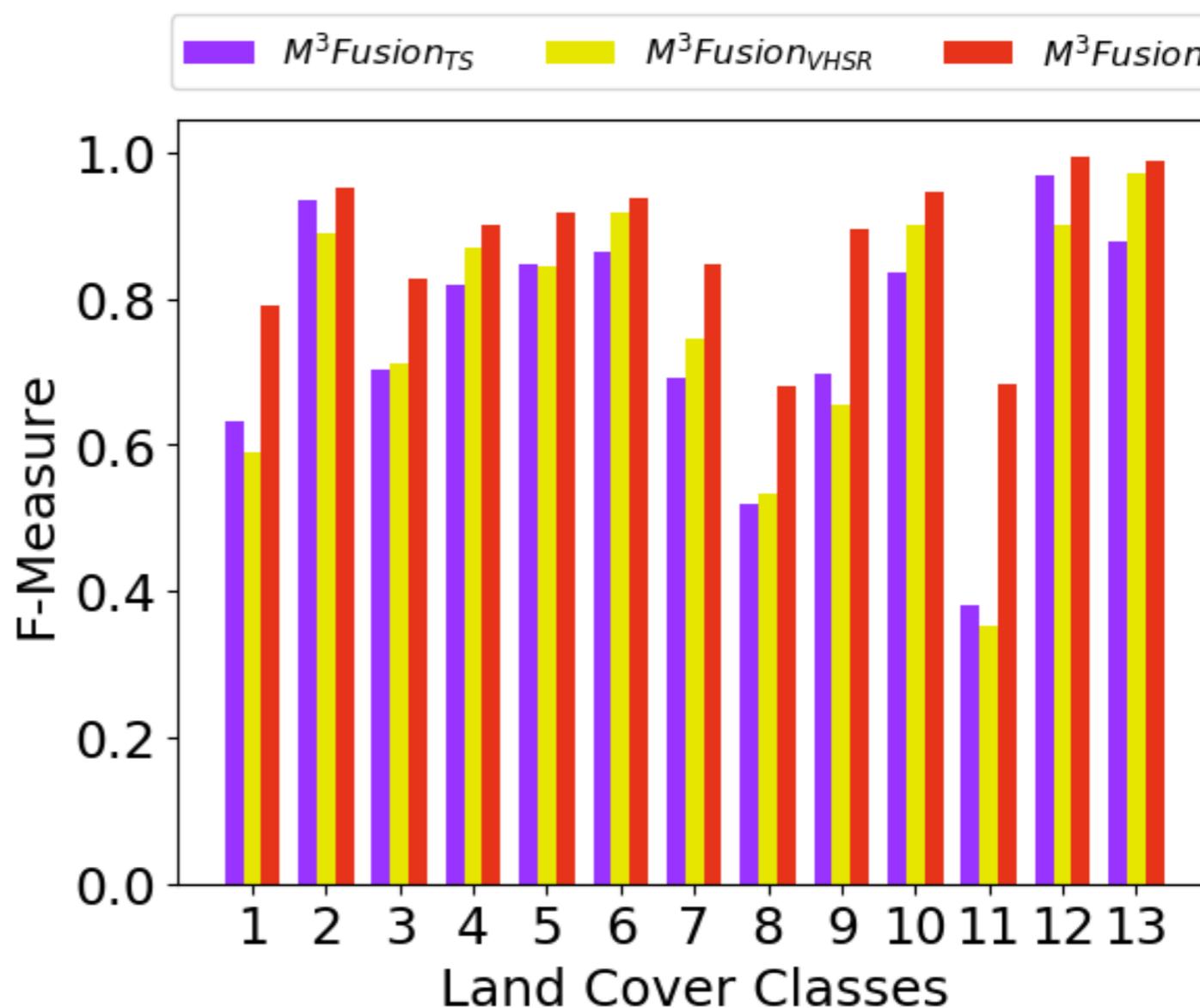
*M<sup>3</sup>Fusion*

	<i>Accuracy</i>	<i>F-Measure</i>	<i>Kappa</i>
<i>RF<sub>TS</sub></i>	0.8543	0.8519	0.8258
<i>M3F<sub>TS</sub></i>	0.8319	0.8325	0.8033
<i>RF<sub>VHSR</sub></i>	0.8237	0.8140	0.7908
<i>M3F<sub>VHSR</sub></i>	0.8369	0.8364	0.8677
RF	0.8716	0.8681	0.8491
M3F	<b>0.9149</b>	<b>0.9148</b>	<b>0.9000</b>

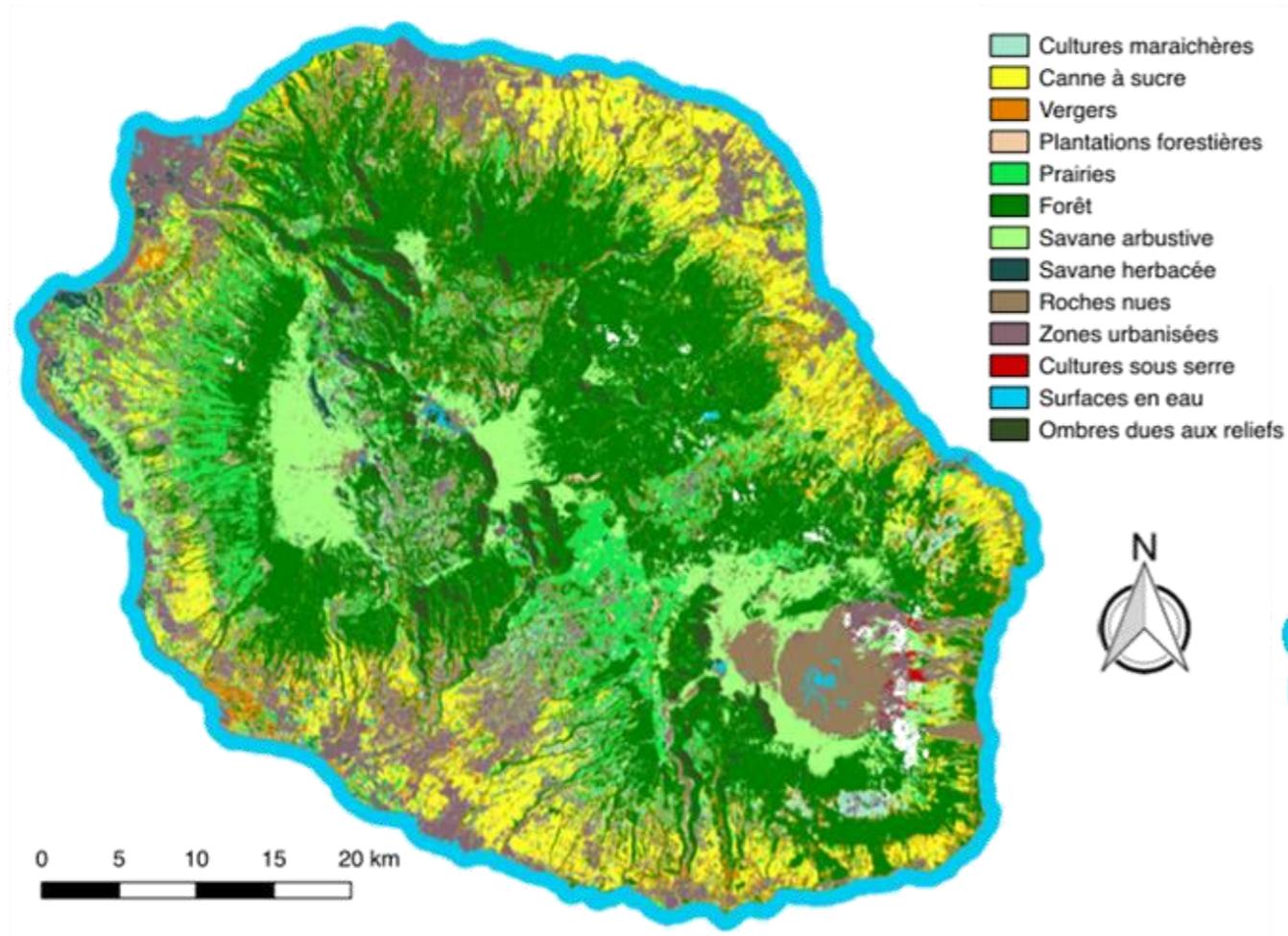
## *M<sup>3</sup>Fusion* : résultats (tests d'ablation)



Class	Label
1	<i>Crop Cultivations</i>
2	<i>Sugar cane</i>
3	<i>Orchards</i>
4	<i>Forest plantations</i>
5	<i>Meadow</i>
6	<i>Forest</i>
7	<i>Shrubby savannah</i>
8	<i>Herbaceous savannah</i>
9	<i>Bare rocks</i>
10	<i>Urbanized areas</i>
11	<i>Greenhouse crops</i>
12	<i>Water Surfaces</i>
13	<i>Shadows</i>



## M<sup>3</sup>Fusion : résultats

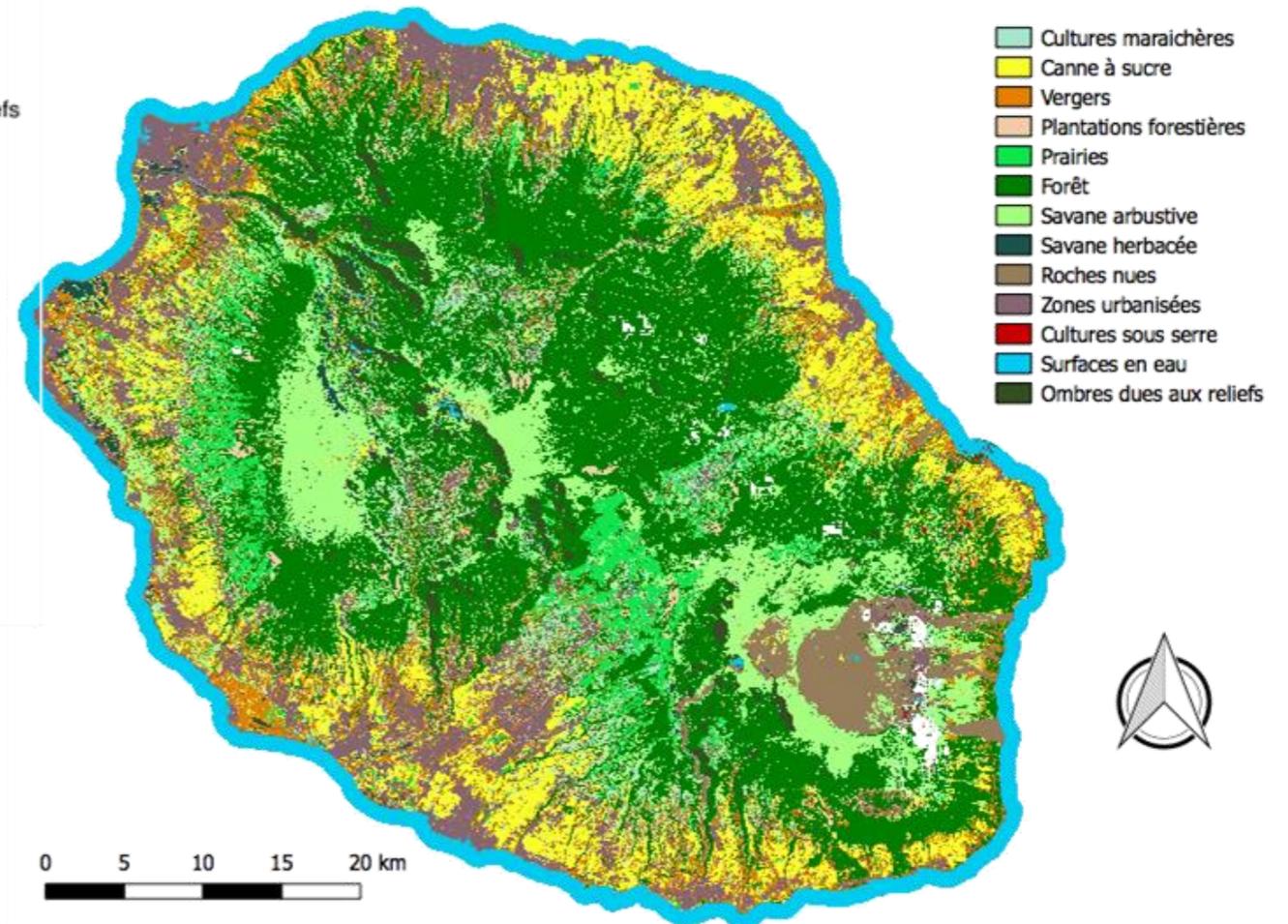


*Random Forest*

Pixels à classer : environ 27M

Temps d'apprentissage M3F : 15h

Temps de production M3F : 9h



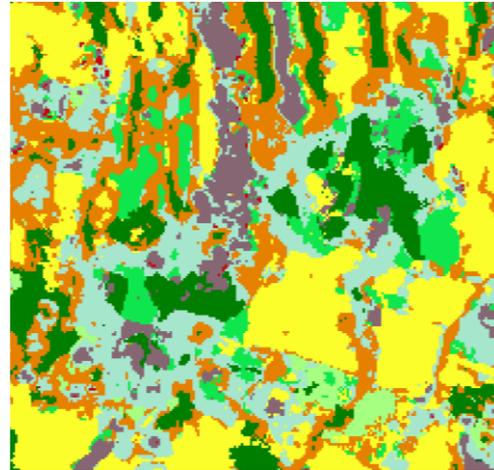
*M<sup>3</sup>Fusion*

## *M<sup>3</sup>Fusion* : résultats

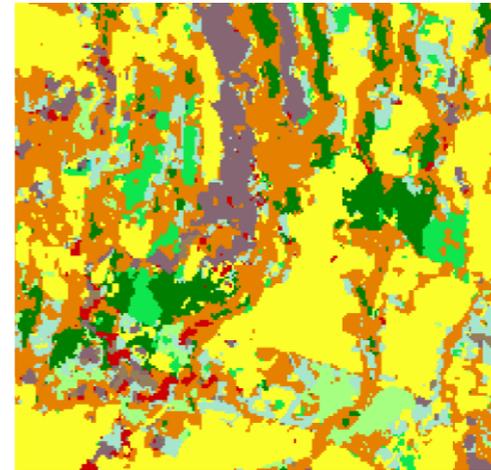
*THRS*



*RF*

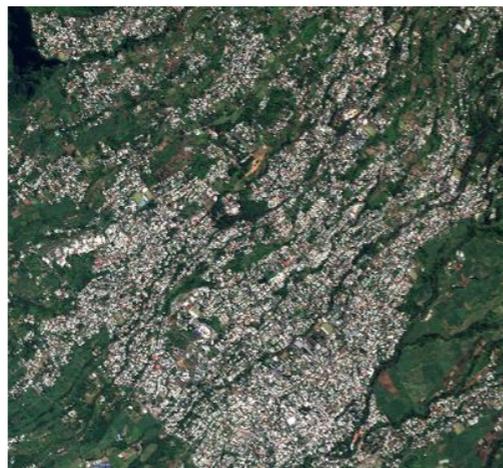


*M3F*

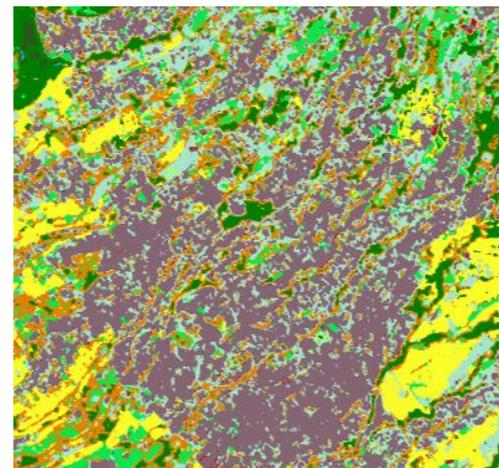


## *Detail sur une zone avec des nuages*

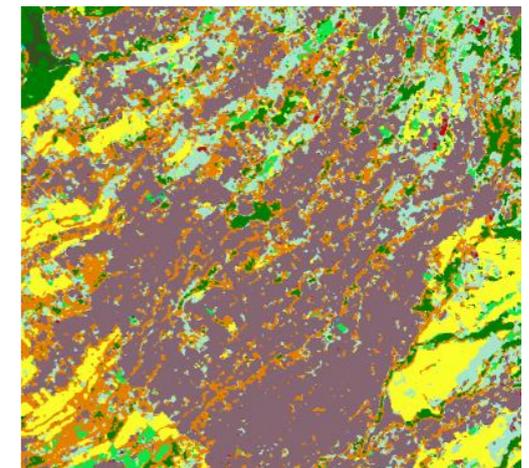
*THRS*



*RF*



*M3F*



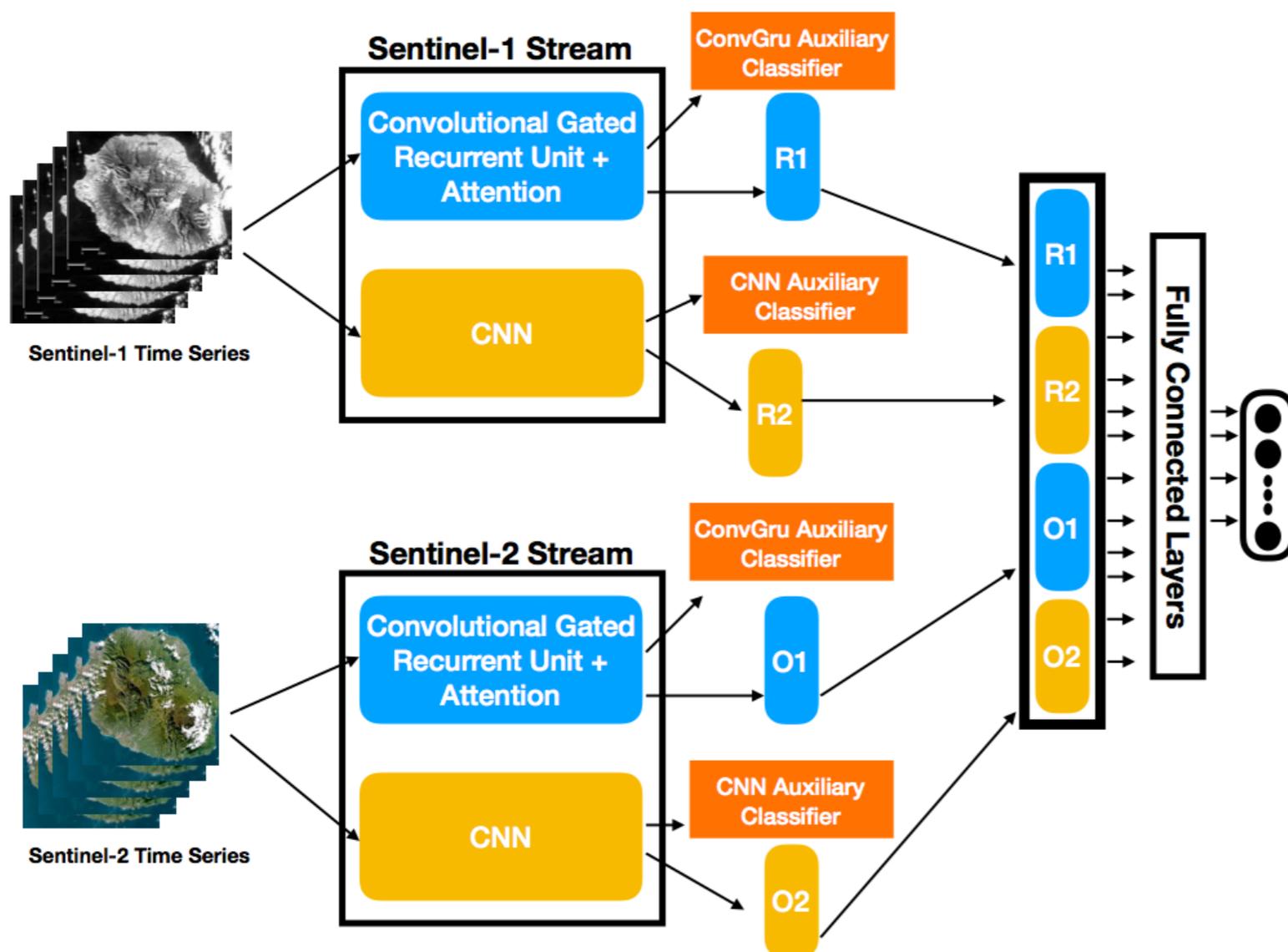
## *Detail sur une zone urbaine*

## **TWINNS** : fusion de séries temporelles d'images optiques et radar via le Deep Learning pour l'occupation du sol [9]

- ❖ Un travail liés à des **activités plus récentes** autour du couplage de séries temporelles **Sentinel-1 et -2** pour améliorer la cartographie de l'occupation des sols
- ❖ **Objectifs** : concevoir une **architecture DL** pouvant mobiliser des **séries temporelles acquises en deux modalités différentes** pour la classification de l'OcSol
  - ❖ Un **défi majeur** dans le contexte de la **fusion multi-capteur**
  - ❖ Exploiter de façon conjointe le **spatial** et le **temporel** dans chaque source
- ❖ **Méthodologie** : encore une architecture à **deux branches opérant en parallèle sur les deux sources disponibles**, basée sur des briques plus complexes (**CNN** et **ConvRNN**)

[9] D. Ienco *et al.*, "Combining Sentinel-1 and Sentinel-2 Satellite Images Time Series for Land Cover Mapping via a Multisource Deep Learning Architecture," ISPRS Journal of Photo. and Rem. Sens., 2019. *Major rev.*

## TWINNS : fusion de séries temporelles d'images optiques et radar via le Deep Learning pour l'occupation du sol [9]



- ❖ Une différence importante : **les deux branches sont du même type**
- ❖ Chacune intègre un double mécanisme :
  - ❖ **ConvRNN : prise en compte du temporelle** tout en s'appuyant sur les contextes spatiaux (par date)
  - ❖ **CNN sur l'ensemble de la série** : « autres » corrélations spatiales entre dates différentes
  - ❖ **Approche DualNET [8]** pour la complémentarité des features

[9] D. Ienco et al., "Combining Sentinel-1 and Sentinel-2 Satellite Images Time Series for Land Cover Mapping via a Multisource Deep Learning Architecture," ISPRS Journal of Photo. and Rem. Sens., 2019. Major rev.

## TWINNS : jeux de données

- **24 images Sentinel-1 (VV/VH)**
- **34 images Sentinel-2 (10 réfl. + 6 indices)**
- Taille : **6656 x 5913** pixels
- Une nomenclature à **13 classes**

- **29 images Sentinel-1 (VV/VH)**
- **23 images Sentinel-2 (10 réfl. + 6 indices)**
- Taille : **5253 x 4797** pixels
- Une nomenclature à **7 classes**

Class	Label	# Objects	# Pixels
1	<i>Crop Cultivations</i>	168	50061
2	<i>Sugar cane</i>	167	50100
3	<i>Orchards</i>	167	50092
4	<i>Forest plantations</i>	67	20100
5	<i>Meadow</i>	167	50100
6	<i>Forest</i>	167	50100
7	<i>Shrubby savannah</i>	173	50263
8	<i>Herbaceous savannah</i>	78	23302
9	<i>Bare rocks</i>	107	31587
10	<i>Urban areas</i>	125	36046
11	<i>Greenhouse crops</i>	49	14387
12	<i>Water Surfaces</i>	96	2711
13	<i>Shadows</i>	38	11400

Class	Label	# Polygons	# Pixels
0	<i>Annual Cropland</i>	671	31 075
1	<i>Fallows</i>	57	1 808
2	<i>Natural Forest</i>	64	15 843
3	<i>Savannah</i>	87	25 156
4	<i>Grassland</i>	142	12 883
5	<i>Rocks</i>	29	852
6	<i>Built up</i>	71	1 096
7	<i>Water</i>	16	1 410

**Un site riche en données de référence**

*Île de la Réunion*

**Un site avec une référence plus faible**

*Koumbia (Burkina Faso)*

## TWINNS : résultats

	<i>F-Measure</i>	<i>Kappa</i>	<i>Accuracy</i>
$RF(S1, S2)$	86.10 ± 0.58	0.8402 ± 0.0065	86.42 ± 0.54
$RF_{LF}(S1, S2)$	87.73 ± 0.58	0.8611 ± 0.0069	88.27 ± 0.59
$2ConvLSTM$	83.21 ± 0.90	0.8031 ± 0.0103	83.17 ± 0.90
TWINNS	89.87 ± 0.65	0.8814 ± 0.0080	89.88 ± 0.69
$RF(TWINNS)$	<b>90.07 ± 1.04</b>	<b>0.8840 ± 0.0124</b>	<b>90.10 ± 1.07</b>

*Site Réunion*

*Site Koumbia*

	<i>F-Measure</i>	<i>Kappa</i>	<i>Accuracy</i>
$RF(S1, S2)$	79.79 ± 5.30	0.7424 ± 0.0694	81.25 ± 5.16
$RF_{LF}(S1, S2)$	84.78 ± 2.36	0.8079 ± 0.0315	86.00 ± 2.35
$2ConvLSTM$	85.73 ± 2.24	0.8165 ± 0.0276	86.48 ± 2.08
TWINNS	<b>86.65 ± 2.50</b>	<b>0.8298 ± 0.0322</b>	<b>87.50 ± 2.44</b>
$RF(TWINNS)$	85.79 ± 2.62	0.8172 ± 0.0351	86.54 ± 2.68

## TWINNS : résultats (tests d'ablation)

	<i>F-Measure</i>	<i>Kappa</i>	<i>Accuracy</i>
TWINNS(S1)	73.22 ± 1.23	0.6926 ± 0.0144	73.89 ± 1.24
TWINNS(S2)	84.29 ± 1.19	0.8159 ± 0.0143	84.26 ± 1.26
FullCNN	87.69 ± 0.85	0.8560 ± 0.0107	87.71 ± 0.92
FullRNN	88.23 ± 1.43	0.8620 ± 0.0169	88.22 ± 1.45
TWINNS <sub>NoAux</sub>	83.92 ± 1.05	0.8109 ± 0.0117	83.84 ± 0.97
TWINNS	<b>89.87 ± 0.65</b>	<b>0.8814 ± 0.0080</b>	<b>89.88 ± 0.69</b>

*Site Réunion*

*Site Koumbia*

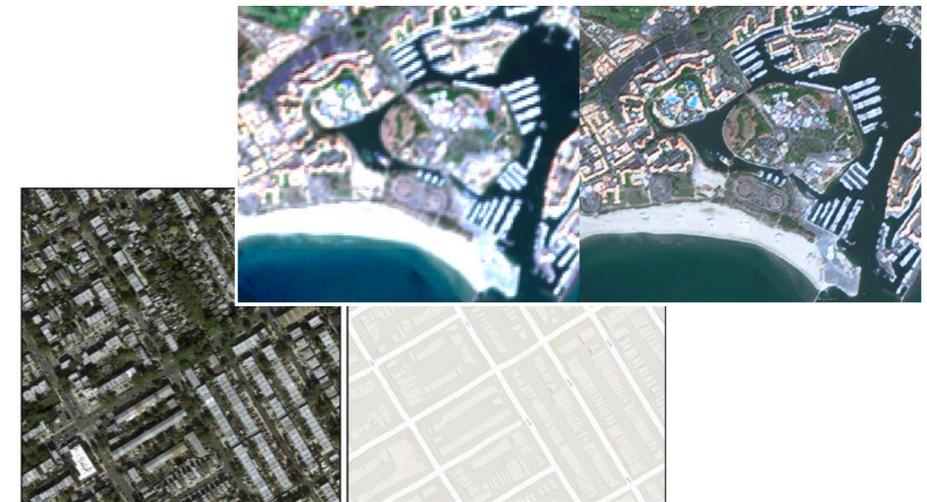
	<i>F-Measure</i>	<i>Kappa</i>	<i>Accuracy</i>
TWINNS(S1)	80.93 ± 2.18	0.7530 ± 0.0283	81.84 ± 2.13
TWINNS(S2)	81.47 ± 4.12	0.7563 ± 0.0556	81.99 ± 4.30
FullCNN	<b>86.81 ± 2.38</b>	<b>0.8303 ± 0.0303</b>	<b>87.51 ± 2.29</b>
FullRNN	85.90 ± 2.72	0.8186 ± 0.0363	86.65 ± 2.75
TWINNS <sub>NoAux</sub>	81.87 ± 4.43	0.7631 ± 0.0599	82.49 ± 4.61
TWINNS	86.65 ± 2.50	0.8298 ± 0.0322	87.50 ± 2.44

## Conclusions et Perspectives

- ❖ L'intérêt vers le **Deep Learning en télédétection** est sans doute très grand
  - ❖ *Un grand nombre de publications en télédétection relèvent de l'application du DL*
  - ❖ *Des résultats prometteurs et de plus en plus « réalistes »*
  
- ❖ Cependant, la plupart de cette littérature se concentre sur l'exploitation directe des outils mises au point en **vision artificielle**
  - ❖ *Les données de télédétection ont leur propres spécificités (échelles, modalités d'acquisition, temporalité, références « éparses », ...)*
  
- ❖ Cela justifie notre intérêt vers les **architectures de Deep Learning « ad-hoc »**
  - ❖ *Adapter les réseaux aux données et pas l'inverse!*
  - ❖ *Une série de travaux montrent les avantages de cette approche et des perspectives plus réalistes de mise en œuvre opérationnelle*
    - ❖ *MultiResoLCC, M<sup>3</sup>Fusion, TWINNS, etc.*

❖ Si on se tient aux sujets « à la mode » en DL, **les perspectives sont très larges**

- ❖ *Generative Adversarial Networks (& co.)*
- ❖ *Spatio/spectro/temporal Domain Adaptation*
- ❖ *Learning in Semi-/Weakly-supervised setting*
- ❖ ...



❖ Un centre d'intérêt à l'**interface avec l'opérationnel et les thématiques**

- ❖ *Des solutions DL dans des conditions de **données manquantes** (nébulosité, acquisitions inexploitable, jeux de données différents entre apprentissage et test, ...)*
- ❖ *Exploitation de la **capacité de transfert des modèles Deep Learning** (ex. Nord/Sud)*
- ❖ *Une longue histoire de modélisation (du signal, physique, interprétative) en TLD → **comment combiner les approches model-driven et data-driven ?***



[Machine and Deep Learning for Earth Observation](https://mdl4eo.irstea.fr)  
<https://mdl4eo.irstea.fr>

Raffaele Gaetano, CIRAD, Research Scientist – Image Processing & Remote Sensing

Roberto Interdonato, CIRAD, Research Scientist - Data Mining and Data Science

Remi Cresson, IRSTEA, Research Engineer - High Performance Computing & Signal processing

Kenji Osé, IRSTEA, Research Engineer - Geomatics and Remote Sensing

Dino Ienco, IRSTEA, Research Scientist - Machine Learning & Data Science

Dinh Ho Tong Minh IRSTEA, Research Scientist - Signal Processing, Radar Imagery

**Merci!**  
**Questions?**