

U-NetMN et SegNetMN: Modèles U-Net et SegNet Modifiés pour la détection des surfaces en eau sur images SAR

Marwane Kzadri
L3IA

Sidi Mohamed Ben Abdellah
Univ Fez, Morocco

Carole Delenne
IUSTI

Aix Marseille Univ., CNRS
Marseille, France

Franco Alberto Cardillo
ILC

National Research Council
Pisa, Italy

Renaud Hostache
Espace Dev

Univ. Montpellier, IRD
Montpellier, France

Nanée Chahinian
HSM

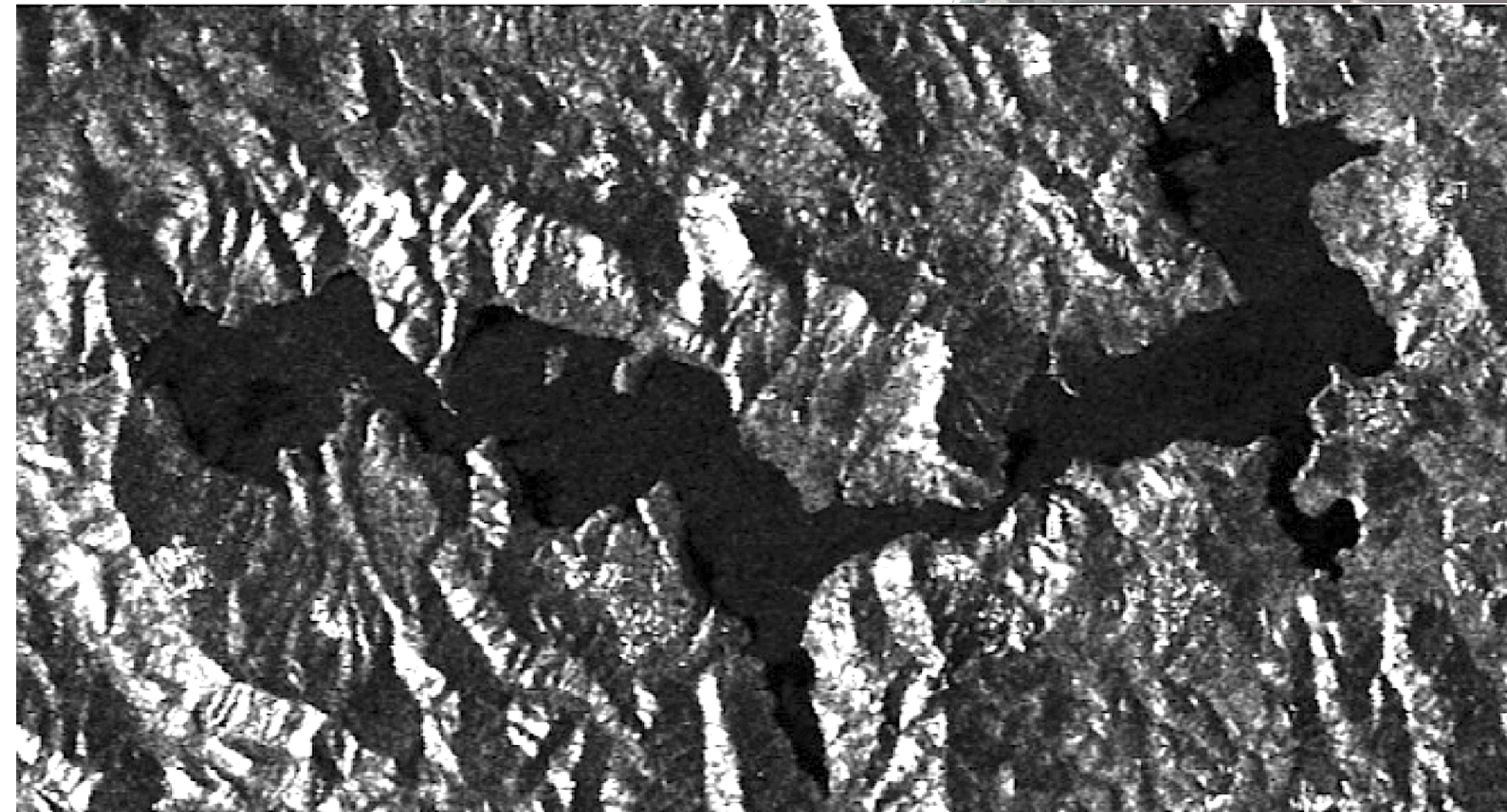
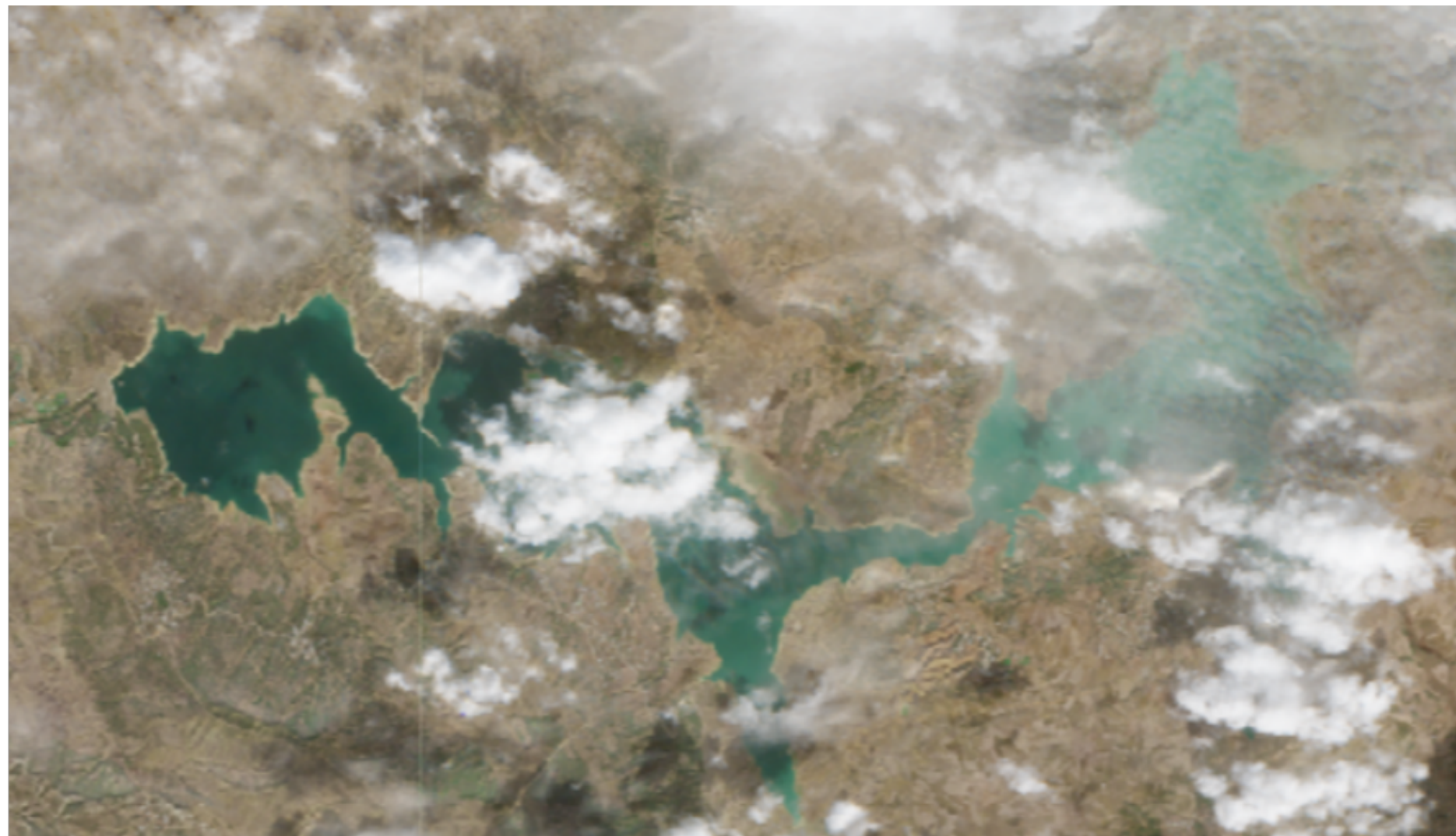
IRD, Univ Montpellier,
CNRS Montpellier, France

Jamal Riffi
L3IA

Sidi Mohamed Ben Abdellah
Univ Fez, Morocco

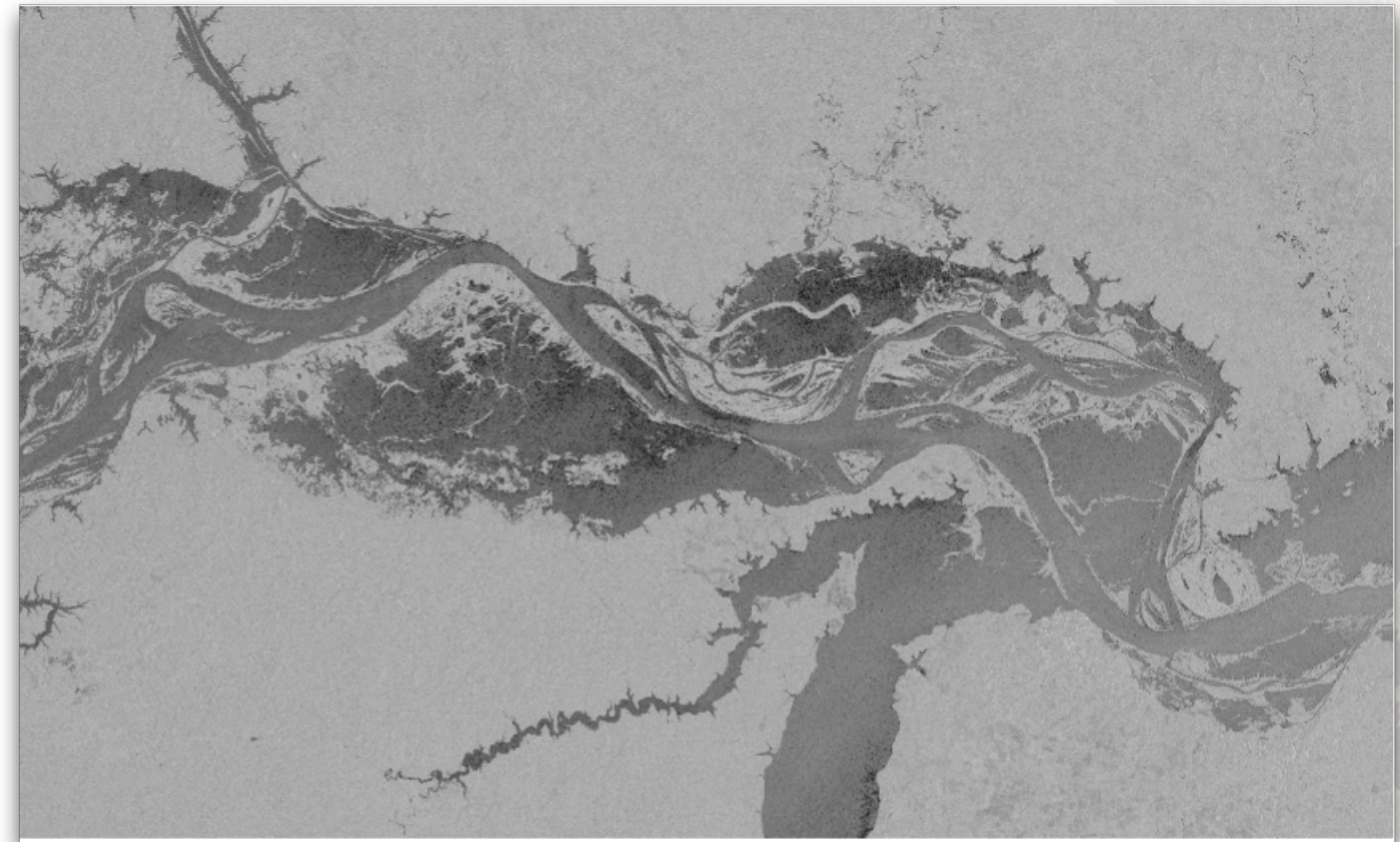
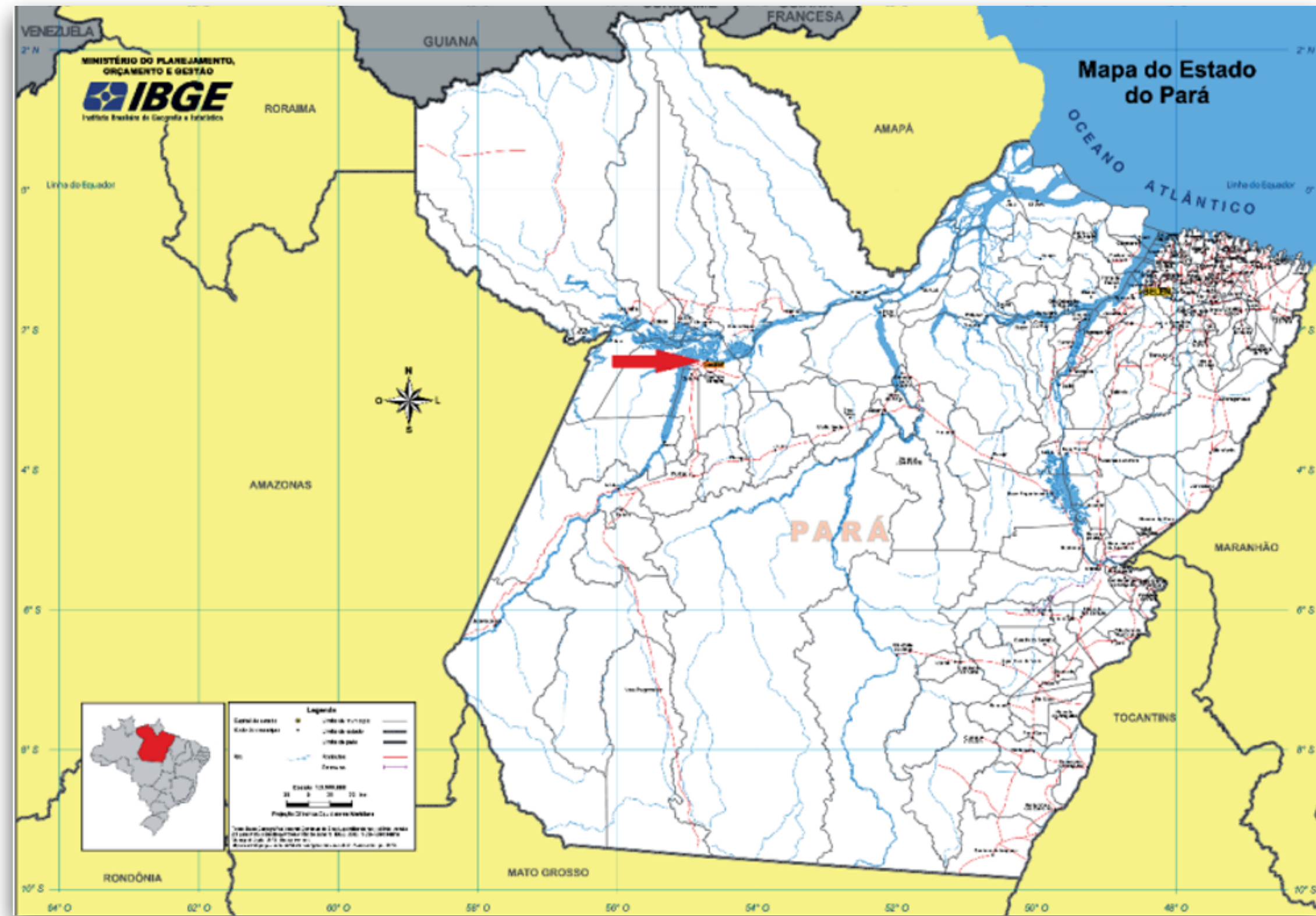
Contexte et Motivation

- ◆ **Détection des plans d'eau:**
tâche essentielle dans la gestion environnementale.
- ◆ **Images SAR / images optiques:**
capacité à traverser les nuages => toutes conditions météo.
- ◆ **Méthodes classiques (seuillage, SVM, forêt aléatoire):**
difficulté de généraliser aux ensembles de données SAR.
- ◆ **Modèles d'apprentissage profond tels que U-Net et SegNet:**
prometteurs, mais convergence lente et variabilité des performances.



Site d'étude

Santarém, Pará, Brésil



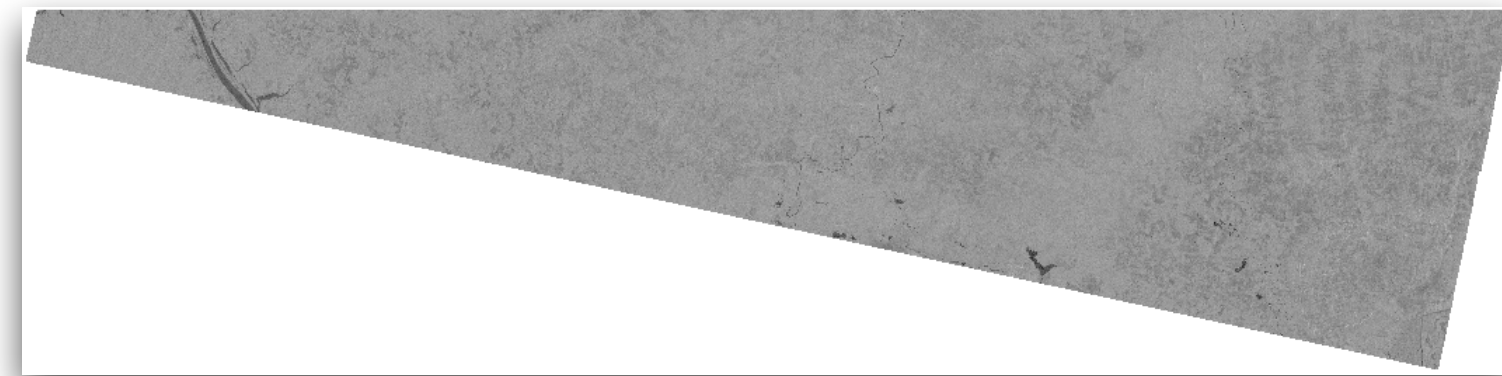
Images Sentinel-1
Résolution 20x20
11112x6706 pixels
~30000km²

Série temporelle ~40 images (de mai 2022 à dec. 2023)

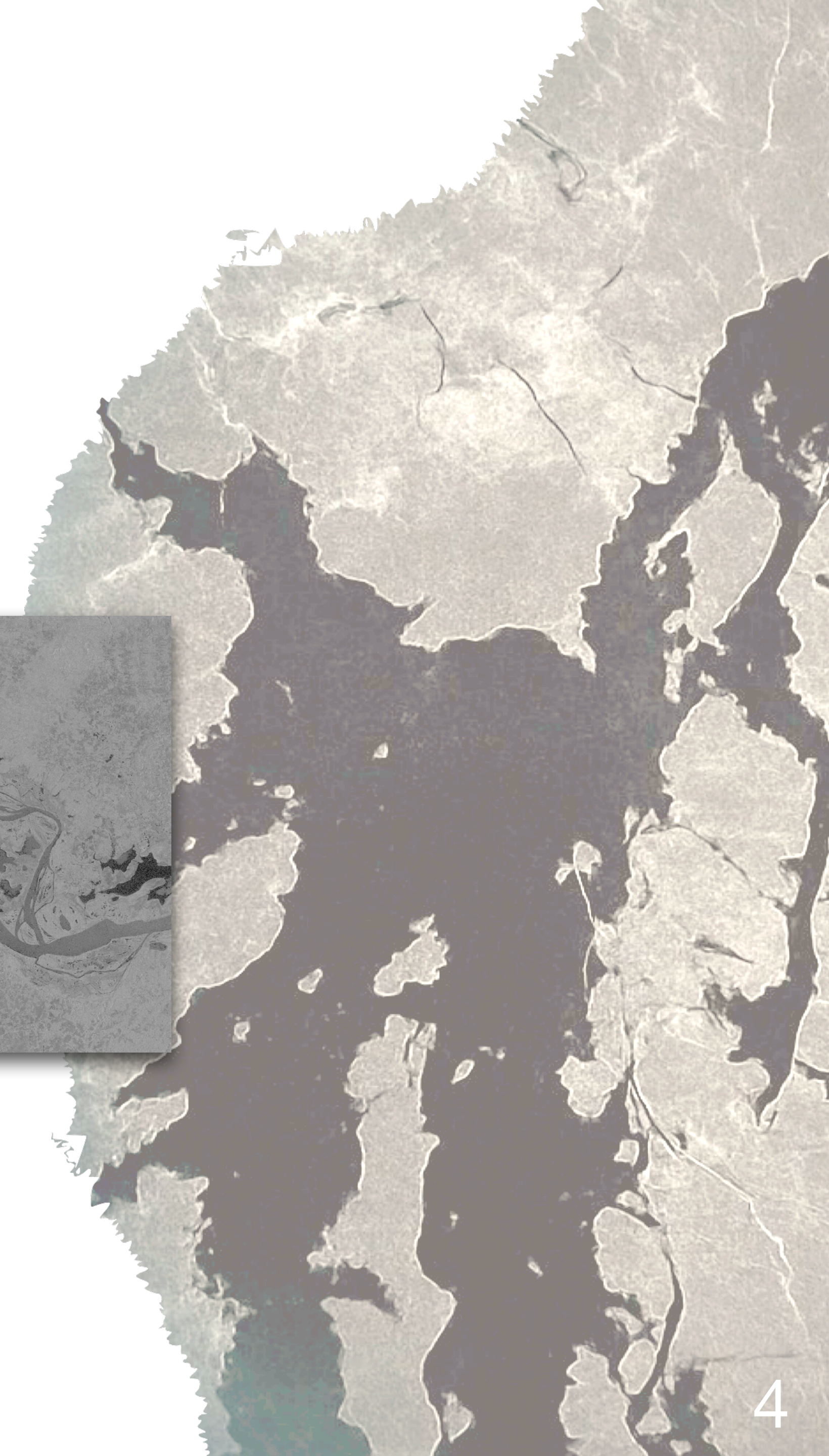
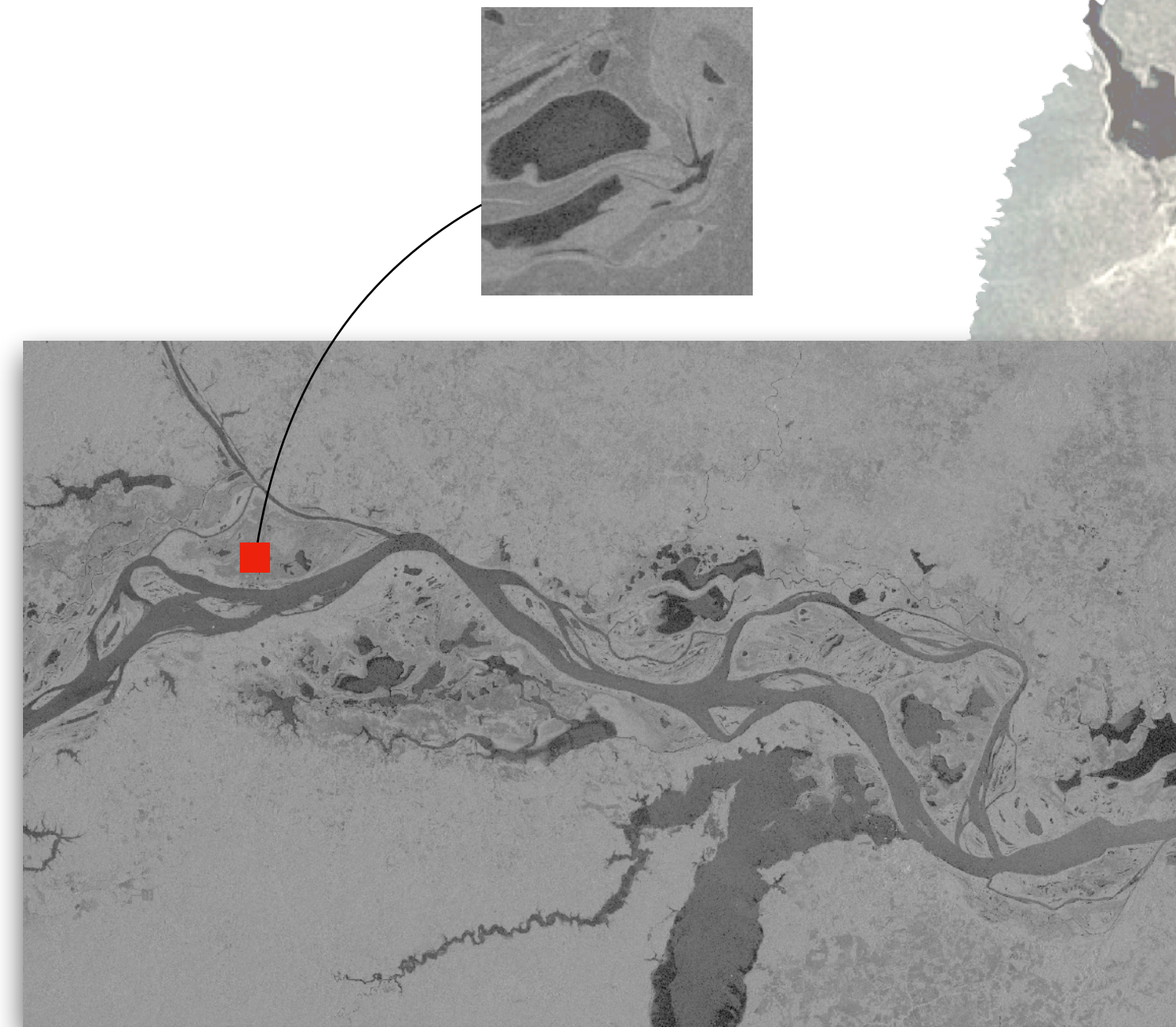
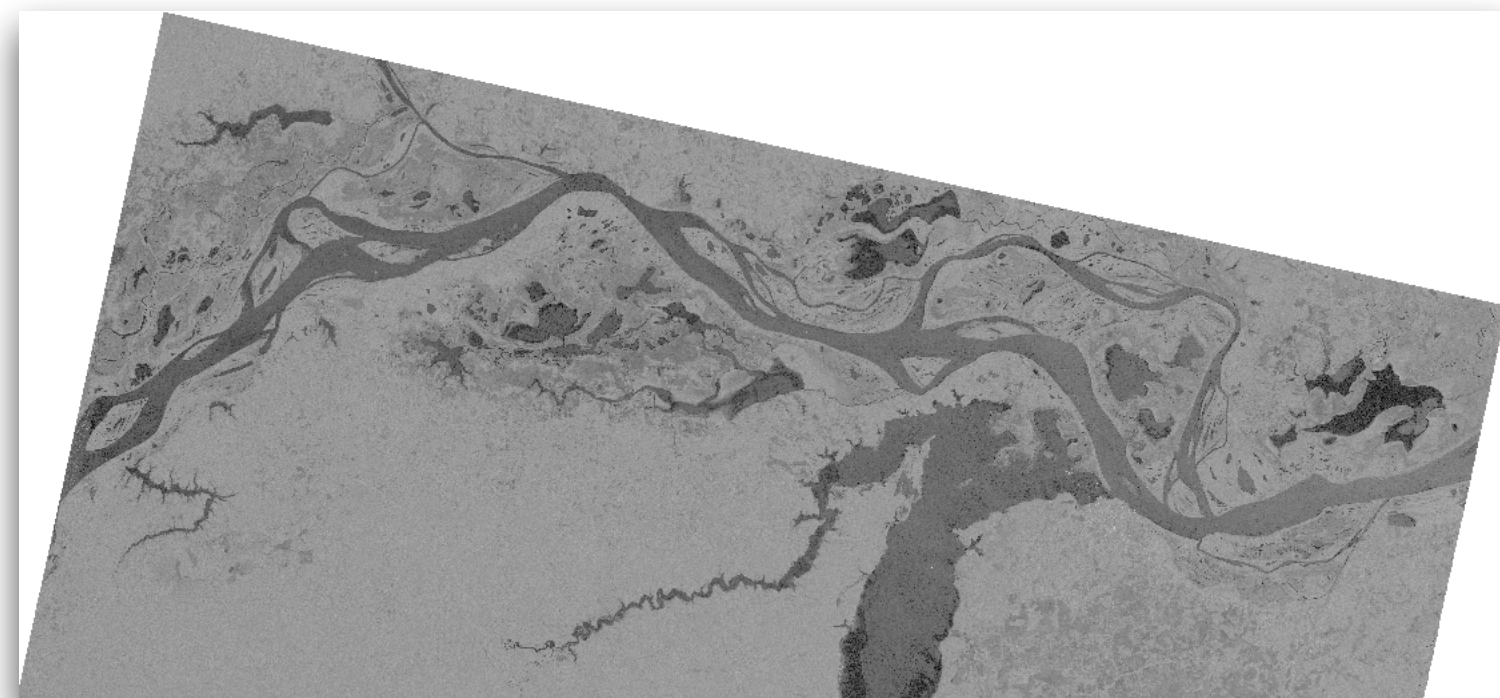
[1] J. B. Oliveira et al., "Flood mapping in the Amazon using SAR images," Rev. Soc. Bras. Med. Trop., vol. 54, 2021.
Available: <https://doi.org/10.1590/0037-8682-0404-2020>

Pré-traitement

- ◆ Zone à cheval sur 2 images => fusion = 11 000 x 6 700 pixels
- ◆ Suppression des nodata
- ◆ Découpage en sous-images de 256x256

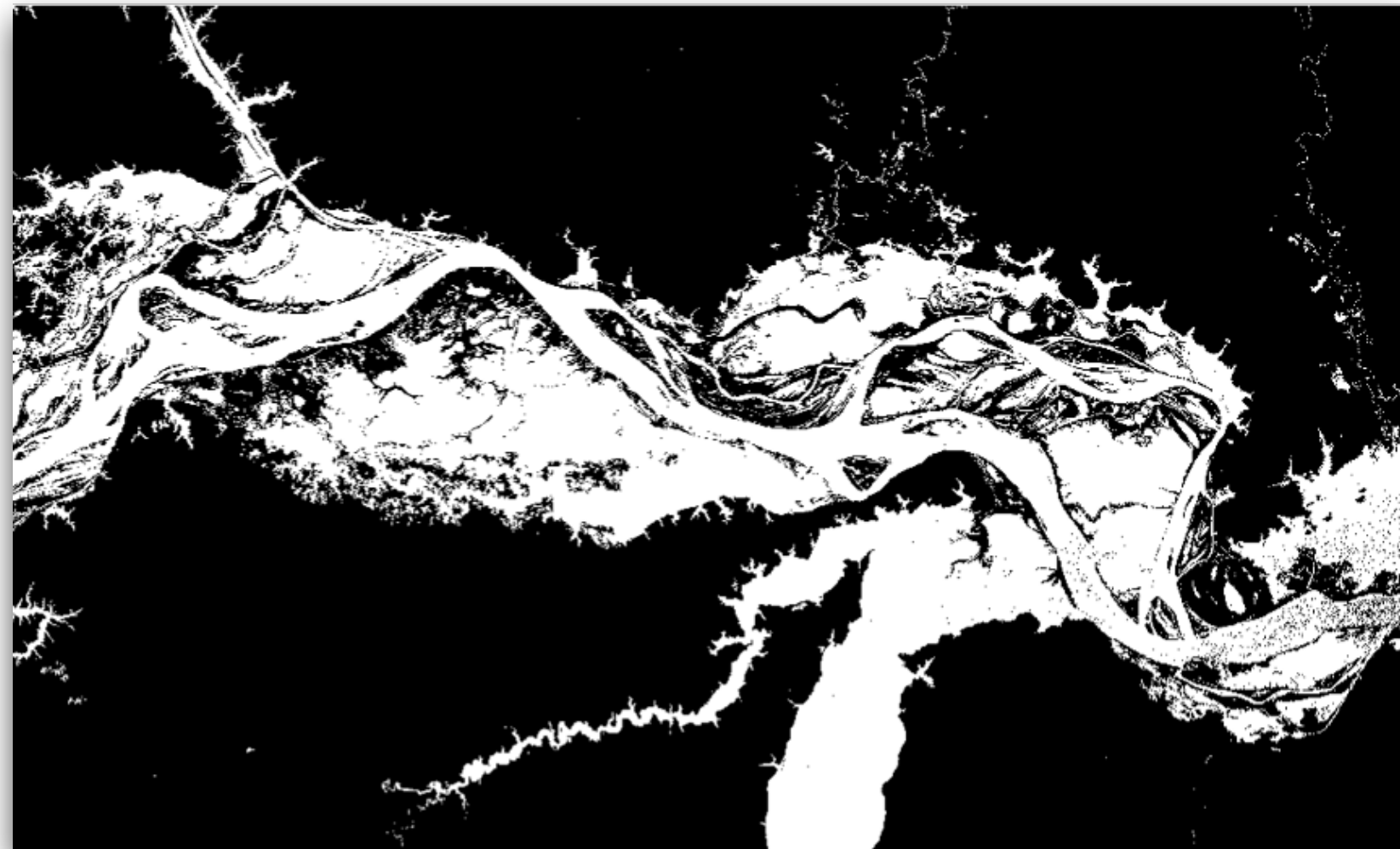


+

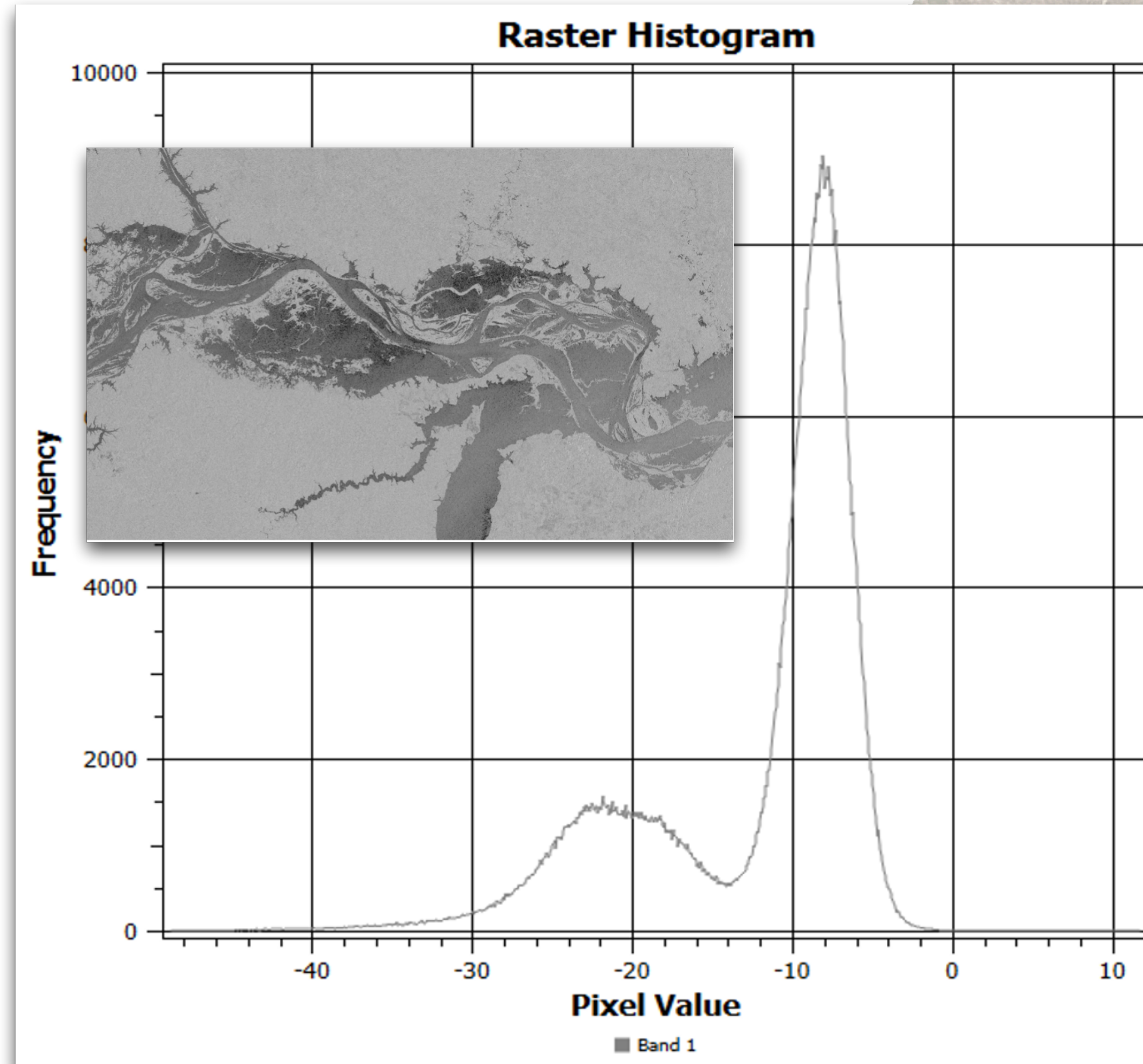


Base de données

- ◆ Histogramme bimodal
- ◆ Masque de référence généré par l'approche HSBA (Hierarchical Split-Based Approach)
- ◆ « Silver ground truth »



M. Chini, R. Hostache, L. Giustarini and P. Matgen. A hierarchical splitbased approach for parametric thresholding of SAR images : Flood inundation as a test case. IEEE Transactions on Geoscience and Remote Sensing, 55(12) :6975–6988, 2017. <https://doi.org/10.1109/TGRS.2017.2737664>



Méthode

U-Net [Ronneberger et al. 2015]

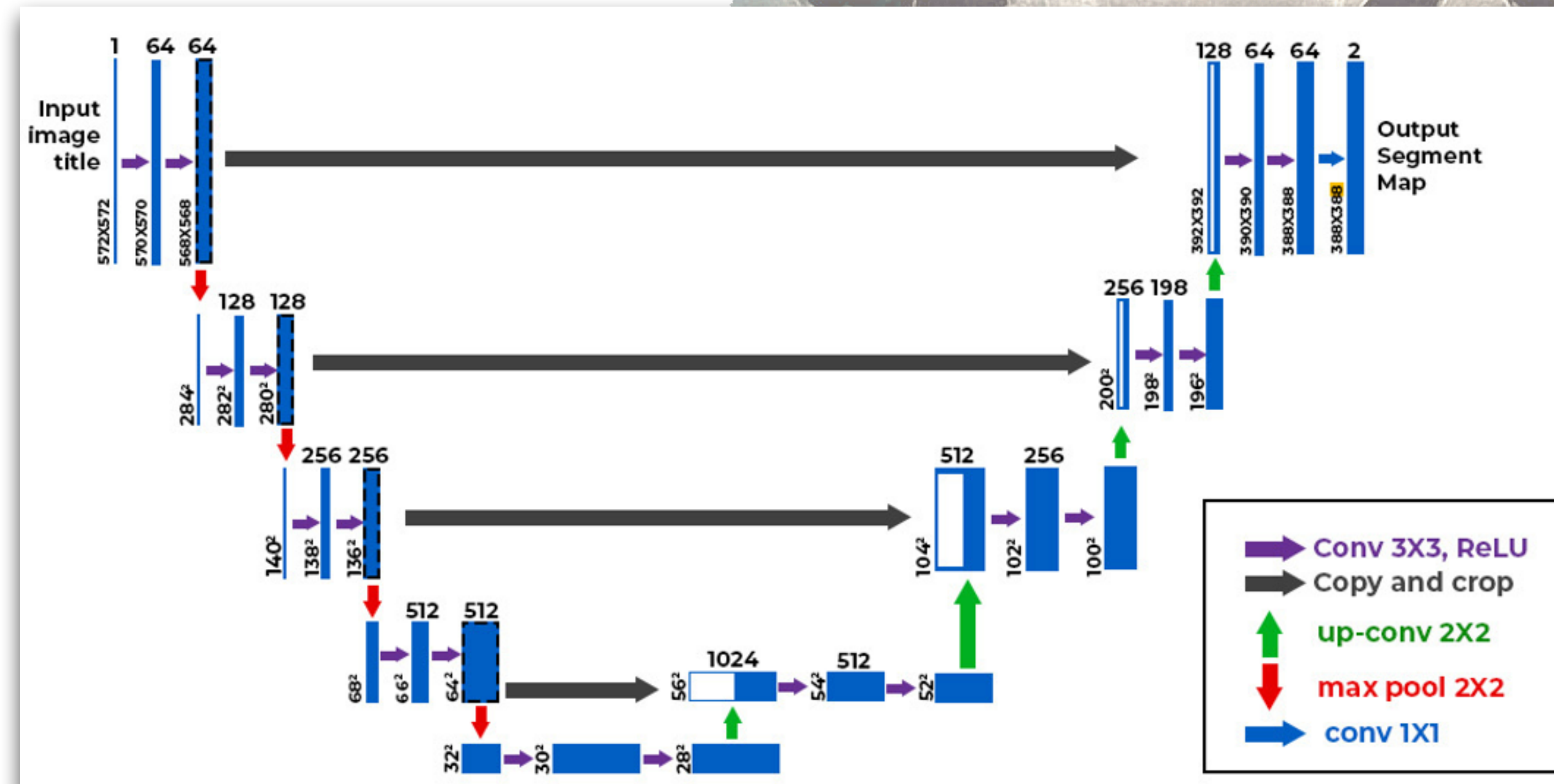
Une des architectures les plus populaires pour la segmentation

◆ Forme en U avec 2 parties principales:

- Encodeur (gauche) extrait les caractéristiques tout en réduisant la résolution spatiale
- Décodeur (droite) restaure les dimensions spatiales pour la classification au niveau des pixels.

◆ Connexions de saut (skip connections) entre les couches dans l'encodeur et le décodeur.

- Permettent au modèle de récupérer des détails fins, pour segmenter des régions petites ou fines comme les rivières ou les zones inondées dans les images SAR.



O. Ronneberger, P. Fischer, and T. Brox,
« U-Net: Convolutional Networks for
Biomedical Image Segmentation » in Proc.
MICCAI 2015, Springer, pp. 234–241.
DOI: 10.1007/978-3-319-24574-4_28

Méthode

SegNet [Badrinarayanan et al. 2017]

- ◆ Structure encodeur-décodeur sans « skip connection »
- ◆ Stocke les indices de max-pooling pendant la phase d'encodage et les utilise pour guider le processus de suréchantillonnage dans le décodeur.
- ◆ Plus légère et rapide que U-Net
- ◆ Parfois moins précise lors de la segmentation de structures fines, telles que les petits plans d'eau dans les images SAR.

V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, Dec. 2017. DOI: <https://doi.org/10.1109/TPAMI.2016.2644615>

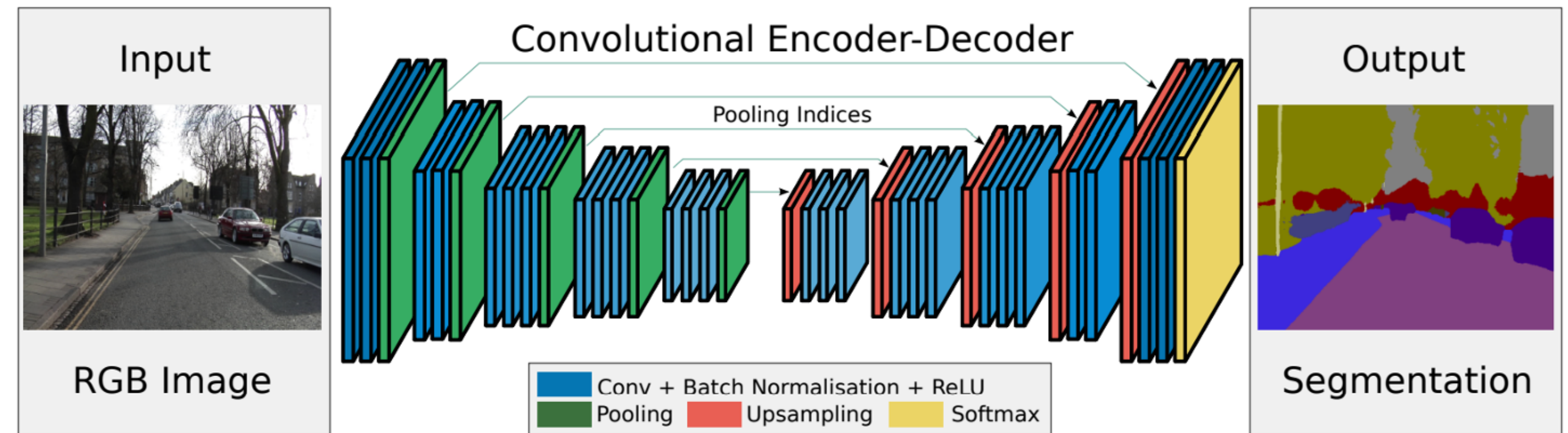
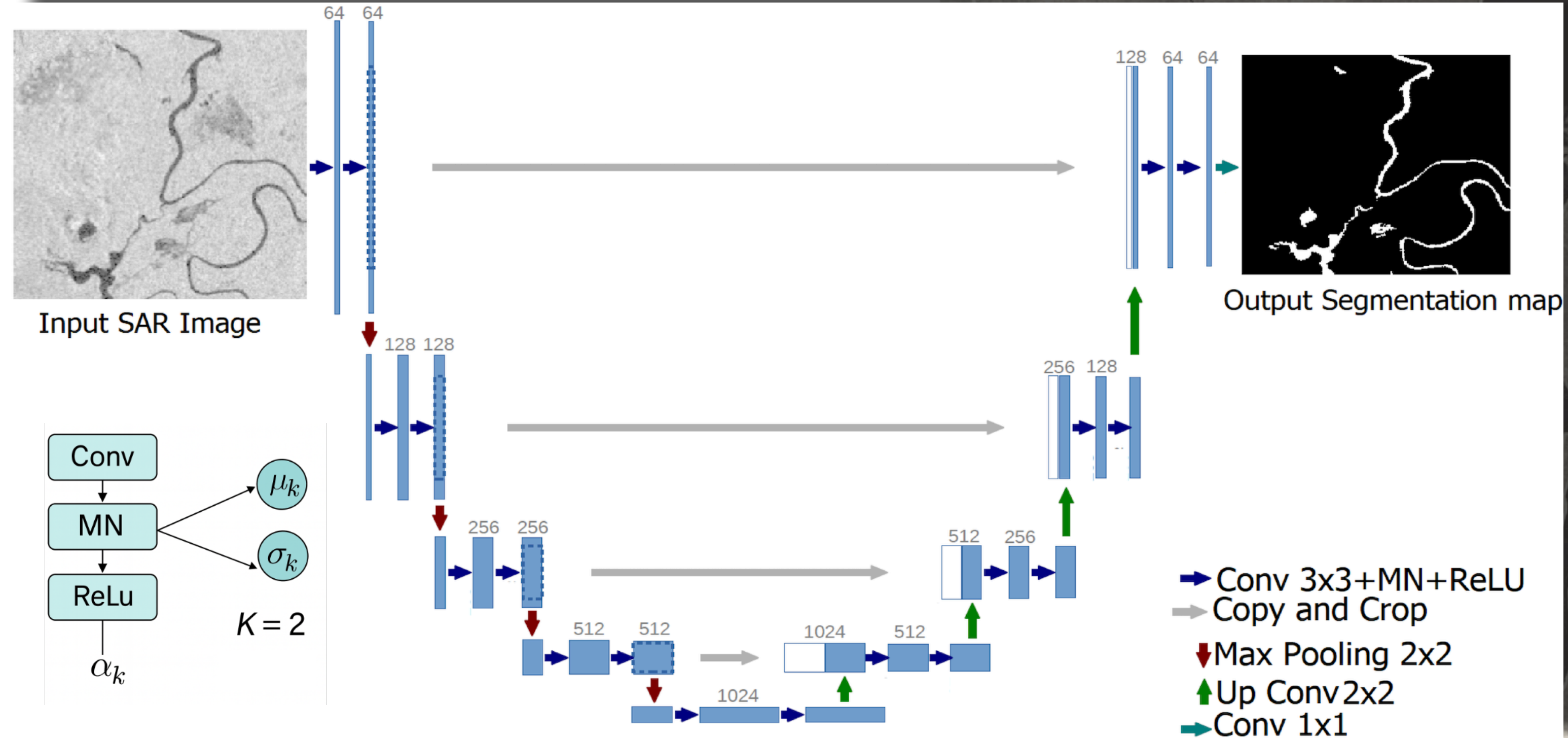


Fig. 2. An illustration of the SegNet architecture. There are no fully connected layers and hence it is only convolutional. A decoder upsamples its input using the transferred pool indices from its encoder to produce a sparse feature map(s). It then performs convolution with a trainable filter bank to densify the feature map. The final decoder output feature maps are fed to a soft-max classifier for pixel-wise classification.

Mode Normalization (MN)

Suppose un histogramme multi-modal

- ◆ Contrairement à la normalisation par lots, qui suppose une distribution gaussienne unique
- ◆ Détecte différents modes et calcule la moyenne et la variance séparément pour chaque mode.
- ◆ Aide le modèle à mieux s'adapter à la distribution des données SAR et améliore les performances de segmentation, en particulier dans les régions hétérogènes.



Protocole

Ensemble de données ~11 000 sous-images

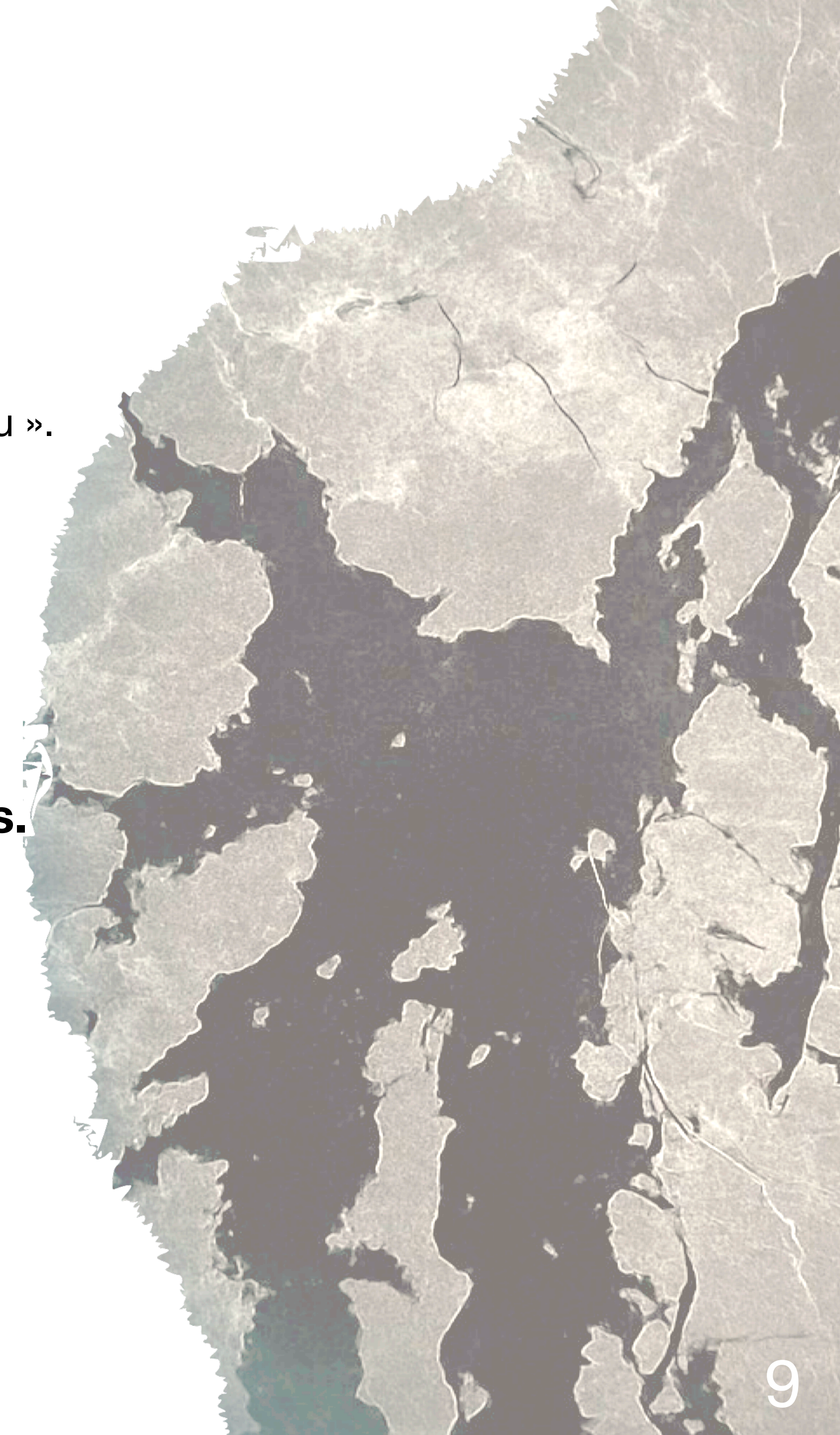
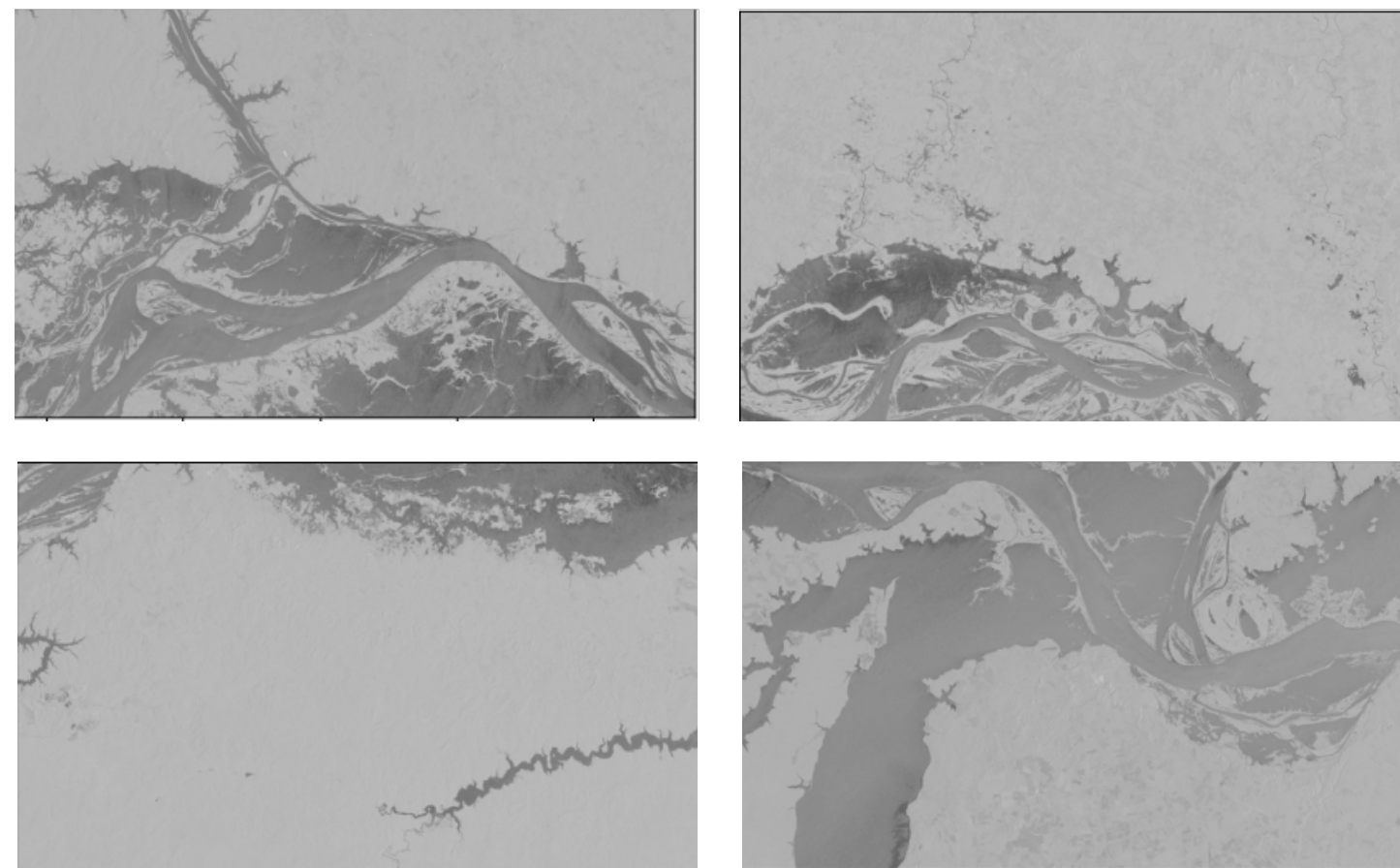
◆ **Division: 70 % entraînement, 10 % validation, 20 % test.**

● Division stratifiée garantissant des ratios équilibrés entre les classes « eau » et « non-eau ».

◆ **Réglage des hyperparamètres à l'aide de la recherche par grille.**

◆ **Arrêt prématuré (early stopping) utilisé pour optimiser l'entraînement.**

◆ **Évaluation : validation croisée à 4 volets basée sur des zones spatiales.**



Résultats

Comparaison U-Net, SegNet et leurs versions Mode Normalization

- ◆ **U-Net plus performant que SegNet**
- ◆ **U-NetMN et SegNetMN légèrement moins performants mais presque 2x plus rapides:**
 - Convergence U-NetMN en 8 époques / 33 pour U-Net
= 1.92x plus rapide
 - Convergence SegNetMN en 12 époques / 32 pour SegNet
=1.85x plus rapide

Modèle	Precision	Recall	Accuracy
U-Net	0.984	0.986	0.991
U-NetMN	0.982	0.979	0.989
SegNet	0.925	0.937	0.961
SegnetMN	0.908	0.914	0.949

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

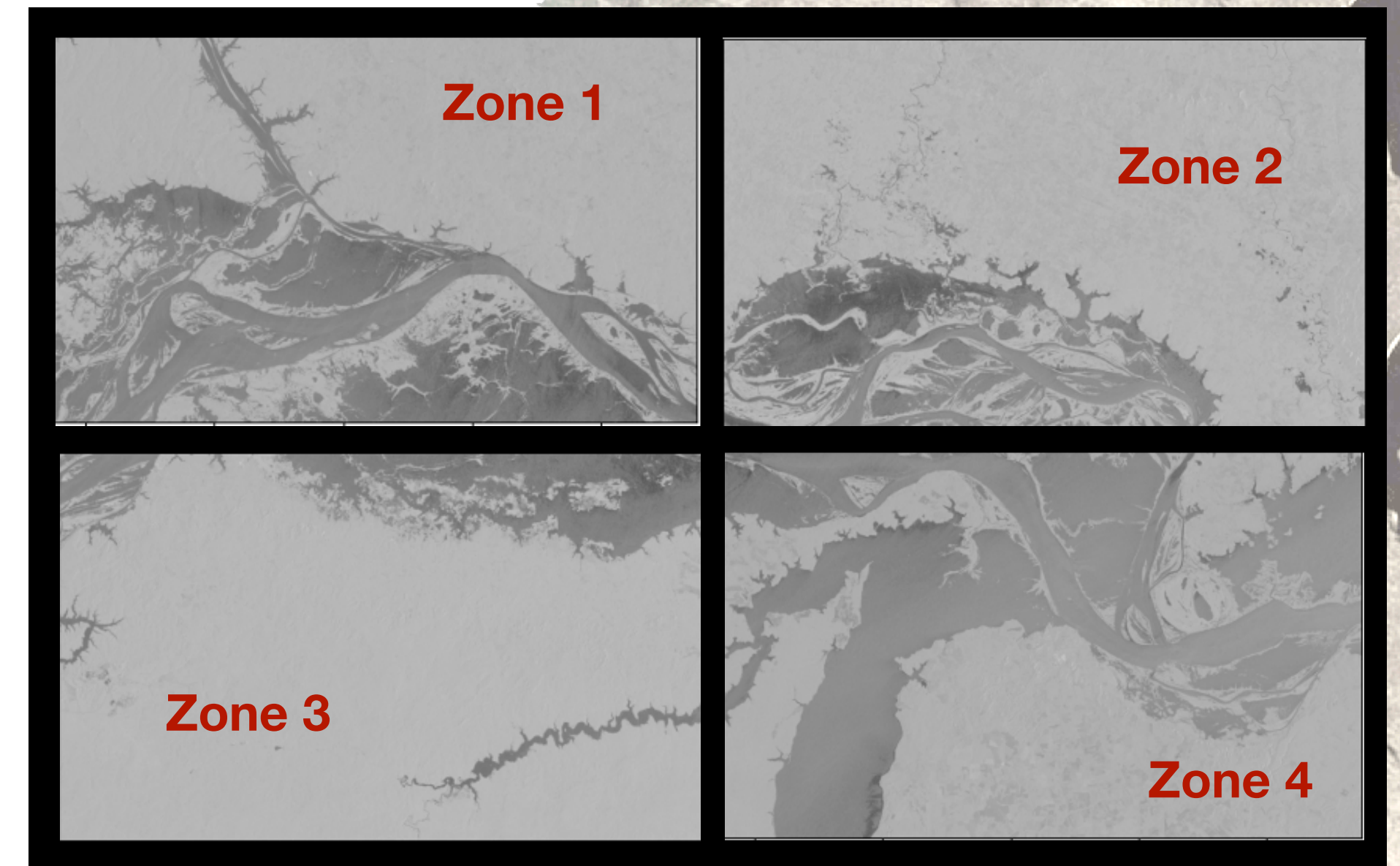
Validation croisée

- ◆ **4-fold validation :**
Images originales coupées en 4 zones de tailles égales
- ◆ **Normalisation de mode**
=> Résultats plus stables sur chaque zones pour SegNetMN

Modèle	Zone 1	Zone 2	Zone 3	Zone 4
U-Net	0.989	0.995	0.993	0.985
U-NetMN	0.984	0.995	0.991	0.985
SegNet	0.915	0.961	0.857	0.954
SegNetMN	0.955	0.971	0.946	0.963

Résultats
identiques

Amélioration



Conclusion et perspectives

- ◆ **Intégration de la normalisation multimodale:**
 - Accélère la convergence
 - Stabilise les résultats sur différentes zones, en particulier où les distributions eau/non-eau sont déséquilibrées
- ◆ **Résultats obtenus sur un « petit » ensemble de données**
 - 11000 sous images mais série temporelle => certaines très ressemblantes
- ◆ **Travaux en cours: généralisation**
 - données multi-site de surfaces en eau
 - images médicales