

Cas d'underfitting dans la segmentation d'images LiDAR

Florian HERGUEZ⁽¹⁾, Thierry BROUARD⁽¹⁾, Jean-Yves RAMEL⁽¹⁾, Clément LAPLAIGE⁽²⁾, Nathanaël LE VOGUER⁽²⁾, Xavier RODIER⁽²⁾

⁽¹⁾ LIFAT, 64 av. Jean Portalis, 37200 Tours

⁽²⁾ UMR CNRS 7324 CITERES/LAT, 33 allée De Lesseps, 37204 Tours

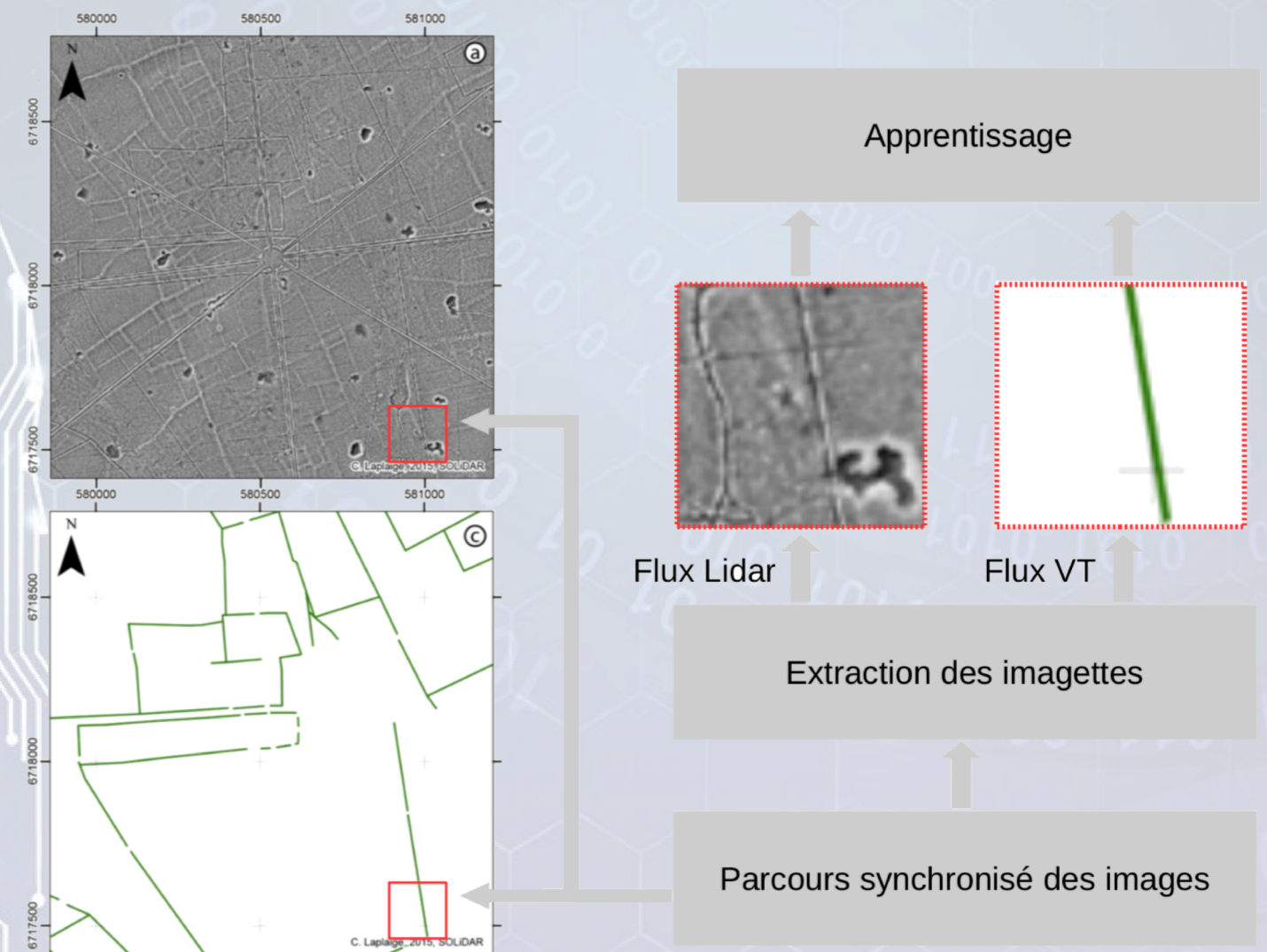
Contexte ► LiDAR est une technologie de télédétection permettant, entre autres, de représenter les micros reliefs d'un terrain tout en éliminant la couverture végétale. Ces images font apparaître des artefacts invisibles sur les prises de vue classiques ou sur le terrain. Ces images sont une source d'information nouvelle pour les archéologues et ils ont aujourd'hui accès à une grande quantité de données de ce type, ce qui a permis de développer des approches utilisant le *deep learning* [1]. Jusqu'à présent on identifiait manuellement les artefacts sur les images. Un précédent travail [2] a mis en place l'architecture nécessaire pour proposer une segmentation automatique utilisant des modèles profonds (*deep learning*).

Entrées ► Il s'agit d'images LiDAR. Chacune est accompagné de sa vérité-terrain, c'est à dire de la segmentation manuelle faite par un archéologue. Le but de l'apprentissage est donc d'entraîner le modèle à donner la même réponse, pour chaque pixel, que la vérité-terrain.

Objectif ► Le but de l'apprentissage est donc d'entraîner le modèle à donner la même réponse, pour chaque pixel, que la vérité-terrain.

Constat ► L'apprentissage du modèle, avant ce travail, semble se dérouler correctement. Cependant l'évaluation des performances montre que le modèle se trompe régulièrement. L'hypothèse est que le manque de caractéristique spécifiques des configurations à rechercher et la forte disproportion entre les pixels « fond » et les pixels « forme » conduit régulièrement au sous apprentissage.

Piste de résolution ► Le prototype actuel utilise la totalité de l'image lors de l'apprentissage. Chaque zone d'intérêt (par ex. un talus, un fossé) étant représentée par très peu de pixels ; se tromper sur ces pixels ne constitue donc qu'une petite erreur, ce qui rend l'apprentissage difficile. L'approche choisie a donc consisté en un fenêtrage systématique de l'image d'entrée, basé sur les informations contenues dans la vérité terrain (VT). Les deux flux, Lidar et VT sont présentés en entrée de la procédure d'apprentissage. Utilisant des fenêtres plus petites, la proportion de pixels de type « forme », utiles pour la segmentation, est plus importante.



Crédits images : [3]

La suite du travail a cherché à évaluer, pour différentes tailles de fenêtrage, et différentes classes à reconnaître (donc différentes configurations de pixels « forme »), la qualité de l'apprentissage d'une part, et la capacité des modèles obtenus à segmenter correctement. Les premiers résultats montrent une faible amélioration de l'underfitting, ce qui nous oriente plutôt vers l'étude d'une nouvelle architecture, plus adaptée. Les travaux se poursuivent donc dans cette direction.

Bibliographie ► [1] Chen, G.; Han, J. A survey on object detection in optical remote sensing images. ISPRS J. Photogramm. Remote Sens. 2016, 117, 11–18. [2] Interactive Deep Learning : Application à la reconnaissance d'éléments archéologiques dans les images LiDAR, V. MAURICE, rapport de PRD, EPU Tours, 2020. [3] SOLiDAR: Diachronie de l'occupation du sol: télédétection LiDAR en forêts de Chambord, Boulogne, Russy et Blois, C. LAPLAIGE et al., rapport de recherche, LAT, 2018.

Remerciements ► Les auteurs tiennent à remercier le RTR DIAMS pour l'aide apportée au financement de ce travail.