

Filtrage et densification d'acquisitions LiDAR par apprentissage profond

Julien Baconat

Superviseurs : Mathieu Bredif
Loic Landrieu

IGN

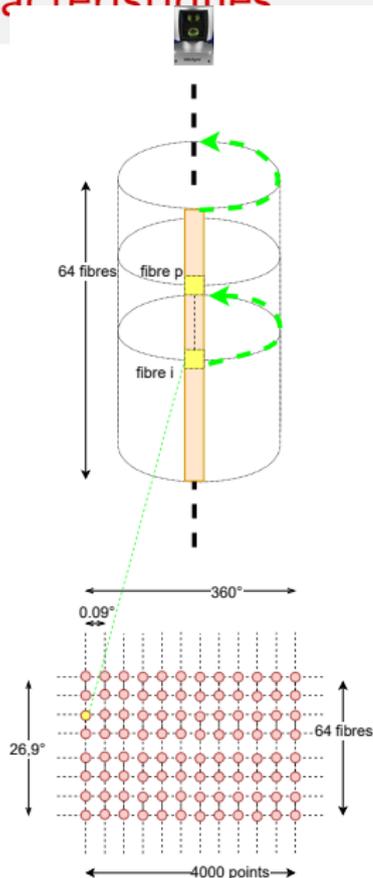
2020

Contexte

- Véhicule d'acquisition mobile IGN
- LiDAR Velodyne HDL-64E
- projet R&D
- Précision métrique à centimétrique

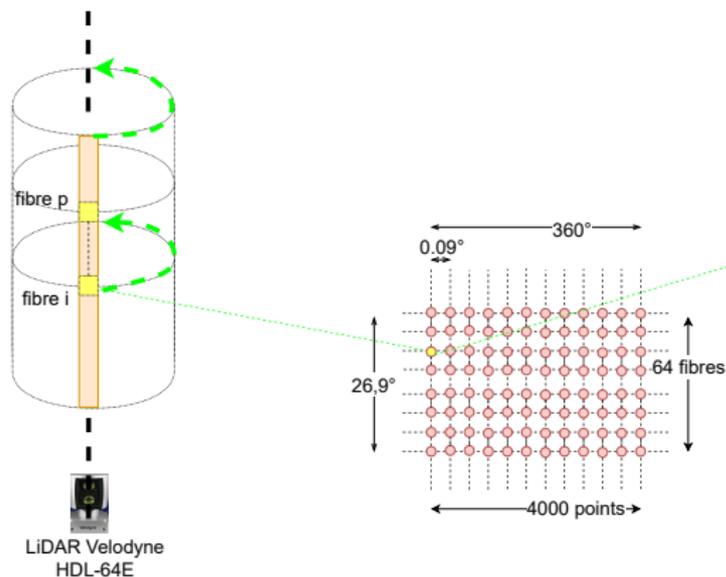


Caractéristiques

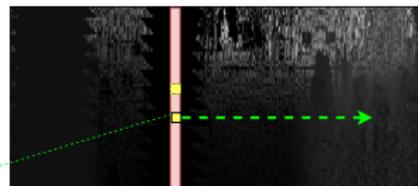


- 64 fibres verticales
- Capacité d'acquisition : 2000 pts /tour et /fibre
- Pas azimuthal : $0.09^\circ \implies$ ligne horizontale de $\frac{360}{0.09} = 4000$ points
- Fréquence de rotation : 10Hz
- $64 \times 2000 \times 10 = 1.3$ millions pts/s

Notion de cartes de données



carte de profondeur



carte du masque de validité



carte de réflectance

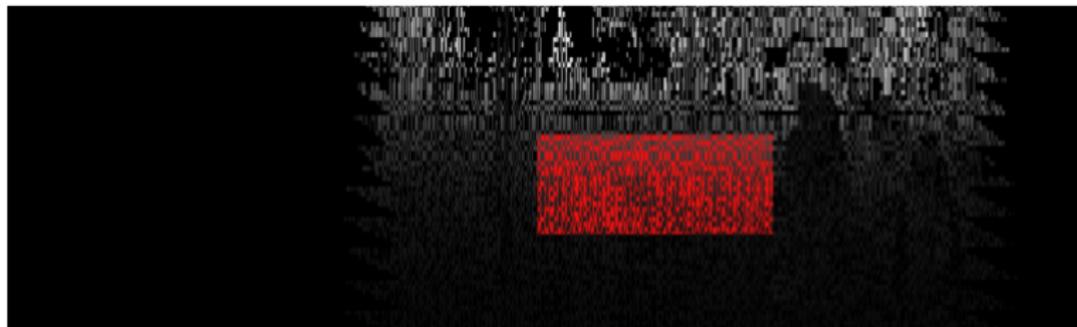


Pixels noirs = valeurs manquantes !

Valeurs manquantes à compléter

2000 points / fibre sur une ligne de 4000 \implies 50% des pixels sont manquants

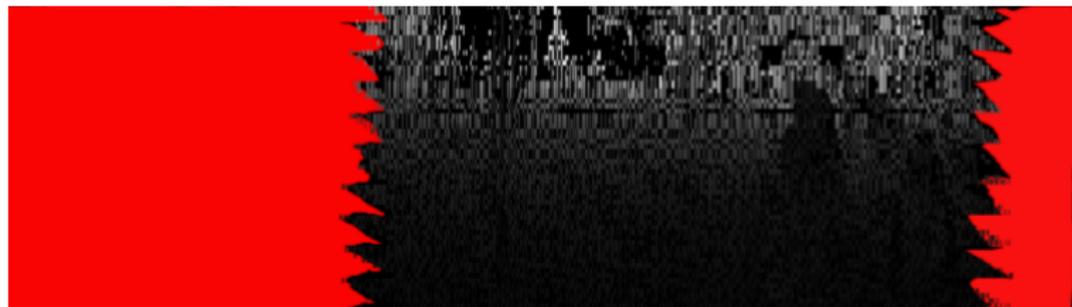
- Cause : capacité d'acquisition du capteur
- Résolution : interpolation
- Difficulté : facile



Valeurs manquantes à compléter

Perte de paquets \implies des parties entières de la carte sont perdues

- Cause : saturation de la bande passante, reconfiguration
- Résolution : reconstitution
- Difficulté : difficile



Valeurs manquantes à inférer

Des pixels de la scène fixe sont occultés

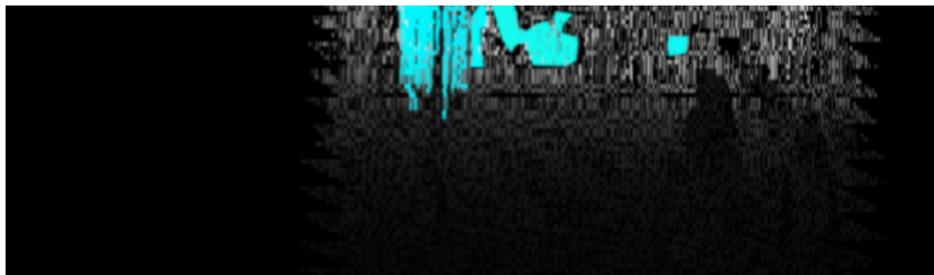
- Cause : objets non pérennes
- Résolution : reconstitution + sémantique latente
- Difficulté : très difficile



Valeurs manquantes ... à ne pas compléter

Certains pixels de la scène ne sont pas observables

- Cause : ciel ou surface non réfléchissantes (vitre)
- Résolution : classification
- Difficulté : difficile



Objectifs

Transformer une carte incomplète en une carte de profondeur dense sans objet mobile

- 1 Densifier à l'exception des pixels non observables
- 2 Supprimer les objets mobiles
- 3 Réduire le bruit d'acquisition

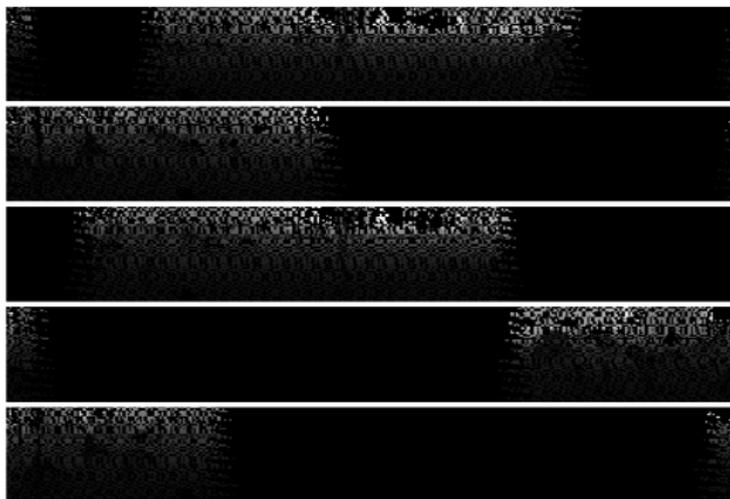
Objectifs

Transformer une carte incomplète en une carte de profondeur dense sans objet mobile

- 1 Densifier à l'exception des pixels non observables
- 2 Supprimer les objets mobiles
- 3 Réduire le bruit d'acquisition

Cas simple : acquisitions statiques sur plusieurs tours

- Combinaison de 30 rotations \implies (1) ✓
- Médiane pour chaque pixel selon l'axe de profondeur \implies (2) ✓ et (3) ✓



combinaison de cartes de profondeur



carte agrégée dense et sans objet mobile

Objectifs

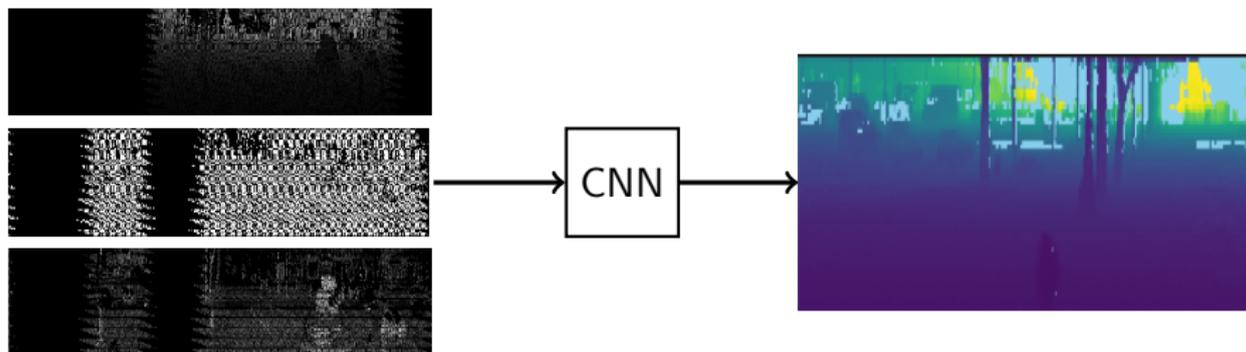
Transformer une carte incomplète en une carte de profondeur dense sans objet mobile

- 1 Densifier à l'exception des pixels non observables
- 2 Supprimer les objets mobiles
- 3 Réduire le bruit d'acquisition

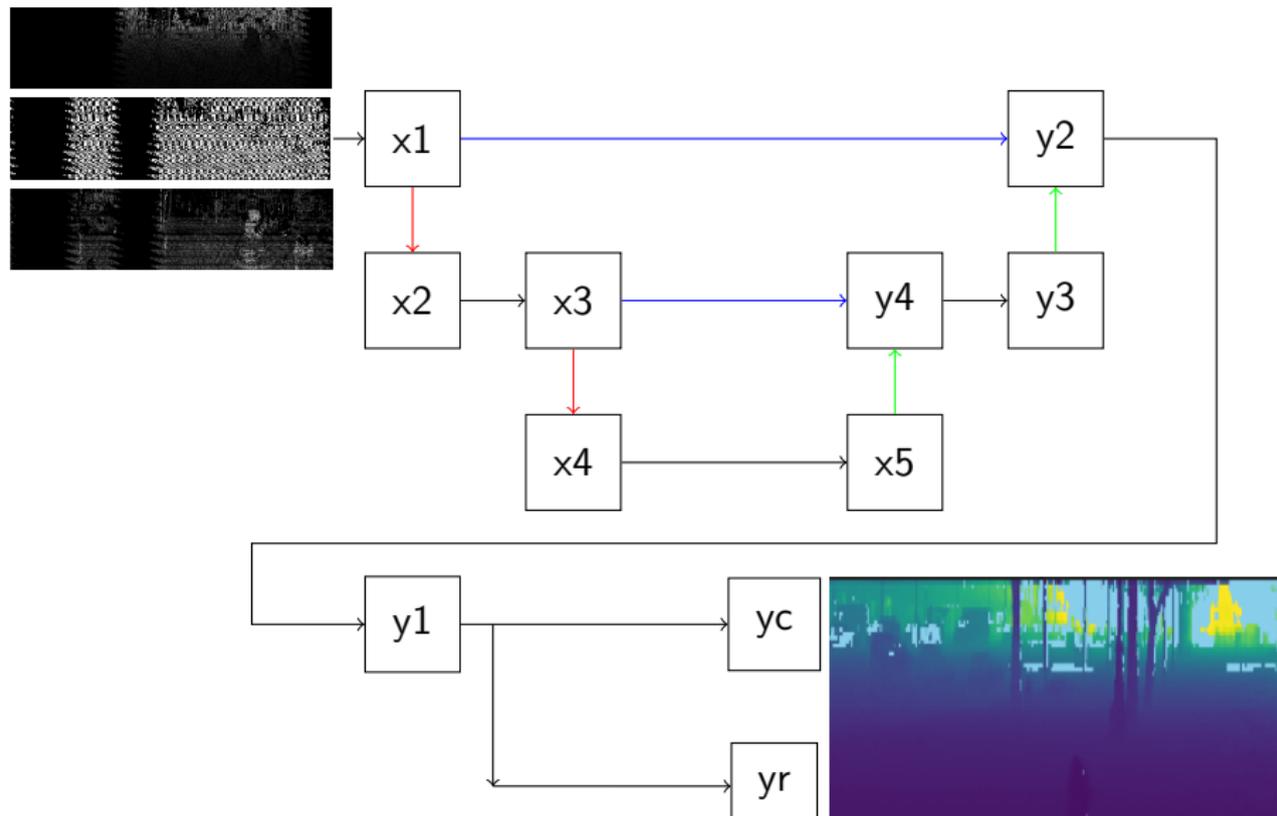
Cas difficile : un tour d'acquisition

- Approche par apprentissage profond
- Cartes agrégées prises comme vérité terrain

Approche par apprentissage profond



U-Net



Fonction de perte

- Multi-tâche
 - 1) detection des pixels non observables (ex: ciel, vitre, ...)
 - 2) évaluation de la profondeur pour les autres pixels
- Entraînement d'une réseau U-Net pour minimiser la fonction de perte

$$\mathcal{L}(p, r) = \lambda \mathcal{L}_{\text{pred}}(p, \hat{p}) + \mathcal{L}_{\text{reg}}(r, \hat{r})$$

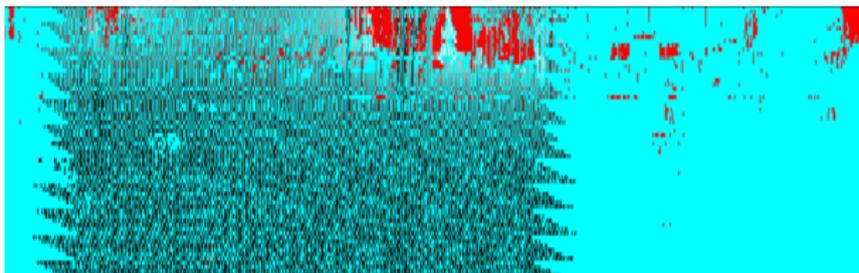
où :

$\mathcal{L}_{\text{pred}}$ = focal loss

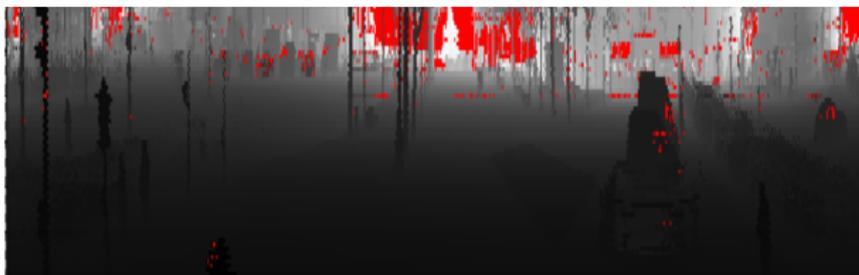
r / \hat{r} : la prédiction de la profondeur / la profondeur pour un pixel dans la vérité terrain pour un pixel

p / \hat{p} : la probabilité qu'un pixel soit non observable / le masque agrégé

Approche par apprentissage profond



Carte de profondeur avec en bleu les valeurs manquantes à compléter(r) et en rouge les valeurs non observables ($\hat{p} = 0$)



Vérité terrain avec en rouge les valeurs manquantes non observables ($\hat{p} = 0$)

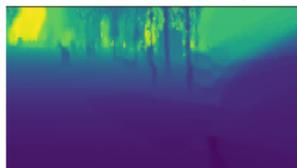
Expérimentations

- Jeu de données de 60 fichiers d'acquisition
 - Entraînement : 46 fichiers (2760 tours)
 - Test et validation chacun sur 12 fichiers (360 tours).
- Entraînement du réseau sur 2000 epochs
- Etude d'ablation
 - Fusion de différentes entrées : +masque ✓ , +reflectance ✗

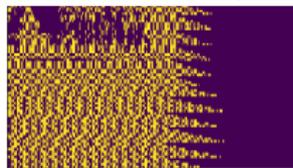
Résultats (1/2)

- Erreur moyenne absolue : $\simeq 3m$
- Accuracy > 97%, IoU : > 88%

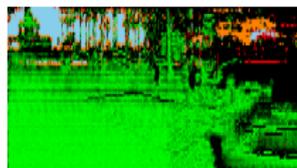
Carte de profondeur (entrée)



Masque de validité (entrée)



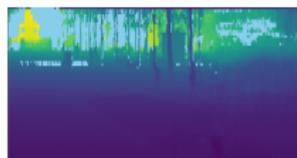
Erreurs



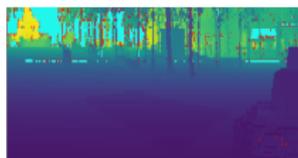
Légende de la carte d'erreur

- en bleu clair : non observables correctement prédits
- en orange : pixels prédits par erreur non observables
- en rouge : pixels non prédits en non observables

Prédiction (sortie)



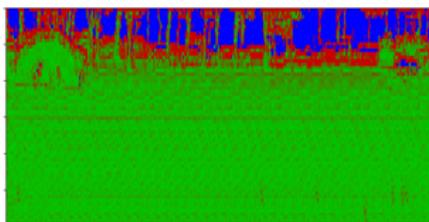
Vérité terrain



Résultats (2/2)

Les fortes erreurs ($> 5m$) sont situées sur les fibres hautes.

Le modèle peine à prédire la profondeur des pixels occultés



carte des erreurs (rouge
: $> 5m$)

carte de profondeur (entrée)



prédiction (sortie)

Conclusions

- Performance
 - Erreur $> 1m$
 - Prédiction des pixels occultés non réalisée
- Causes suspectées
 - Phénomène de sur-apprentissage
 - Manque de données utiles : avec perte de paquets + objets mobiles

Perspectives

- Augmenter le jeu de données
 - par symétrie
 - simulant des pertes de paquets
- Diminuer le nombre de paramètres
- Introduire une régularisation
 - regularisation L2
- Implémenter une couche résiduelle en fin de réseau
- Utiliser un GAN pour obtenir des résultats plus réalistes



O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.



Maximilian Jaritz, Raoul de Charette, Emilie Wirbel, Xavier Perrotton, Fawzi Nashashibi. Sparse and Dense Data with CNNs: Depth Completion and Semantic Segmentation. 3DV 2018 – 6th international conference on 3D Vision. 2018



Pierre Biasutti, Jean-François Aujol, Mathieu Brédif, Aurélie Bugeau. Range-Image: Incorporating sensor topology for LiDAR point cloud processing. Photogrammetric engineering and remote sensing, Asprs American Society for Photogrammetry, 2018, 84 (6), pp.367–375.

Merci pour votre attention

