

Reconstruction de Carte de Profondeur Dense à partir d'un LiDAR 2D et d'une Caméra RGB

Présentation d'une méthode innovante pour la perception robotique, combinant capteurs économiques et apprentissage profond.

OBJECTIF

Reconstruire une Carte de Profondeur Dense en Temps Réel

L'objectif central de ce projet est de développer une méthode robuste et efficace pour surmonter les limitations des capteurs 2D.



Caméra RGB

Image couleur pour le contexte visuel détaillé.

LiDAR 2D

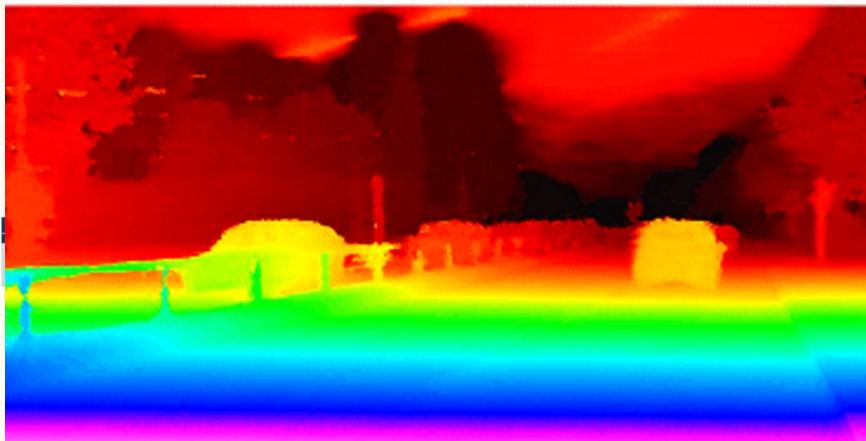
Mesure de profondeur sparse et économique.

Carte de Profondeur Dense

Information 3D complète et détaillée.

La méthode doit garantir une exécution embarquée en temps réel, essentielle pour les applications de robotique et de véhicules autonomes.

Le Dataset KITTI : Un Référentiel pour l'Apprentissage



Le dataset KITTI est un standard de l'industrie pour la recherche en perception autonome, offrant une richesse de données calibrées.

- *Images RGB Frontales* : Fournissent des informations visuelles cruciales pour la perception.
- *Nuages de Points LiDAR 3D* : Utilisés comme vérité terrain pour l'apprentissage supervisé.
- *Matrices de Calibration Précises* : Permettent une cohérence géométrique parfaite entre les capteurs.

Cette approche permet une simulation réaliste d'un LiDAR 2D à partir du LiDAR 3D, garantissant une supervision de haute qualité pour l'entraînement du modèle.

Simulation du LiDAR 2D et Vérité Terrain

Afin de créer un environnement d'apprentissage contrôlé et réaliste, nous simulons les données LiDAR 2D et la vérité terrain dense.

Input: Simulation LiDAR 2D (Ta ligne)



Simulation LiDAR 2D

Une tranche horizontale du nuage de points LiDAR 3D est extraite pour simuler une mesure de LiDAR 2D.

Image RGB - Sample 6732



Projection Caméra

Les points LiDAR sont projetés dans le repère caméra, associant une profondeur à chaque pixel.

Vérité Terrain (Ground Truth KITTI)



Qénération Vérité Terrain

Une carte de profondeur dense est générée, plafonnant les distances pour éviter les valeurs aberrantes et garantir la précision.

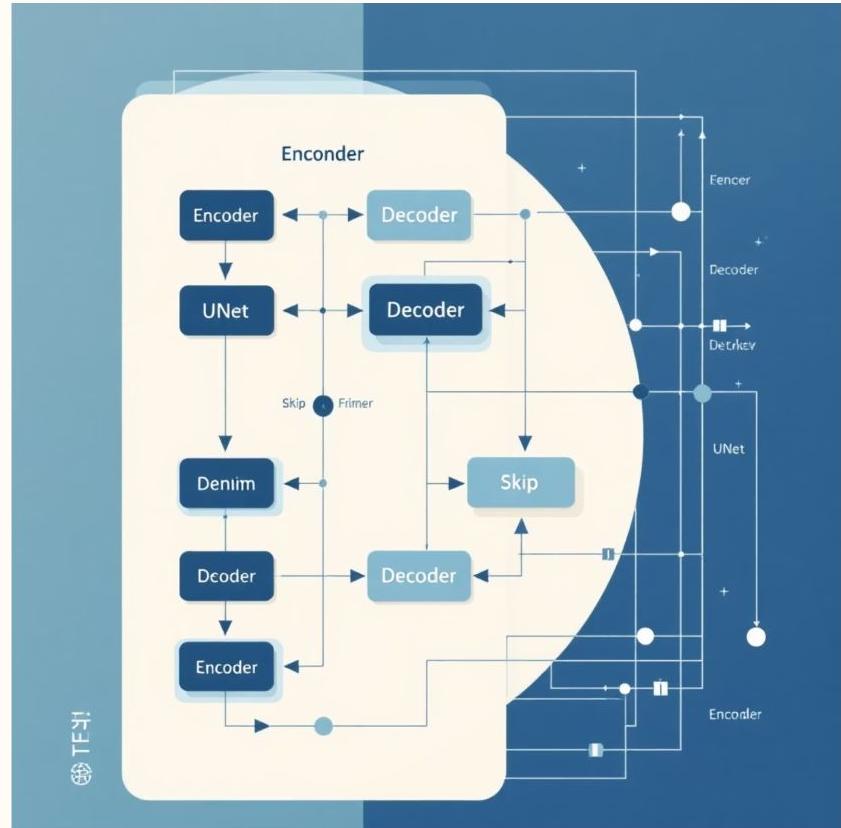
Formulation du Problème et Architecture UNet

Le problème est formulé comme une tâche de régression dense pixel par pixel, où le réseau apprend à inférer la profondeur.

L'entrée du réseau est un tenseur multi-canaux:

- *Image RGB (3 canaux) : Apporte le contexte visuel détaillé.*
- *Profondeur Sparse Simulée (1 canal) : Fournit des indices de profondeur initiaux.*

La sortie est une carte de profondeur dense normalisée, reflétant la structure 3D de la scène.



Fonction de Perte Adaptée et Entraînement



La vérité terrain n'étant pas dense partout en raison des projections LiDAR, une fonction de perte adaptée est cruciale.

Problème des Pixels Non Renseignés

Certains pixels n'ont pas de profondeur valide dans la vérité terrain.

Solution : Masque

Un masque est appliqué pour calculer l'erreur uniquement sur les pixels disposant d'une profondeur valide.

Perte Utilisée

L'erreur quadratique moyenne masquée (masked MSE) est employée pour une évaluation précise du modèle.

L'entraînement utilise l'optimiseur Adam avec un taux d'apprentissage modéré et un batch réduit, le tout sur une séparation entraînement/validation du dat

Visualisation Qualitative et Performance Globale

Image RGB - Sample 6733



Input: Simulation LiDAR 2D (Ta ligne)



Prédiction: Depth Map générée



Vérité Terrain (Ground Truth KITTI)

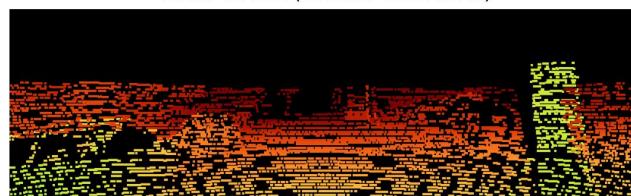


Image RGB - Sample 6734



Input: Simulation LiDAR 2D (Ta ligne)



Prédiction: Depth Map générée



Vérité Terrain (Ground Truth KITTI)

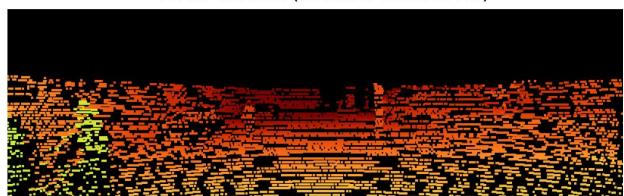


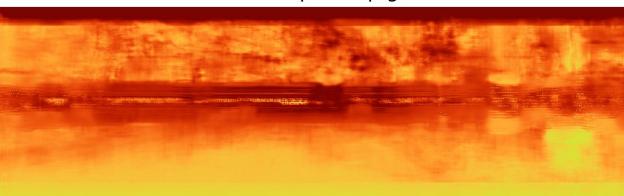
Image RGB - Sample 6736



Input: Simulation LiDAR 2D (Ta ligne)



Prédiction: Depth Map générée



Vérité Terrain (Ground Truth KITTI)



Optimisation pour l'Exécution Embarquée

Pour garantir l'intégration du modèle dans des systèmes embarqués, des étapes d'optimisation spécifiques sont mises en œuvre.



Export du Modèle

Conversion du modèle PyTorch en format ONNX, un standard ouvert pour la représentation de modèles d'apprentissage profond.



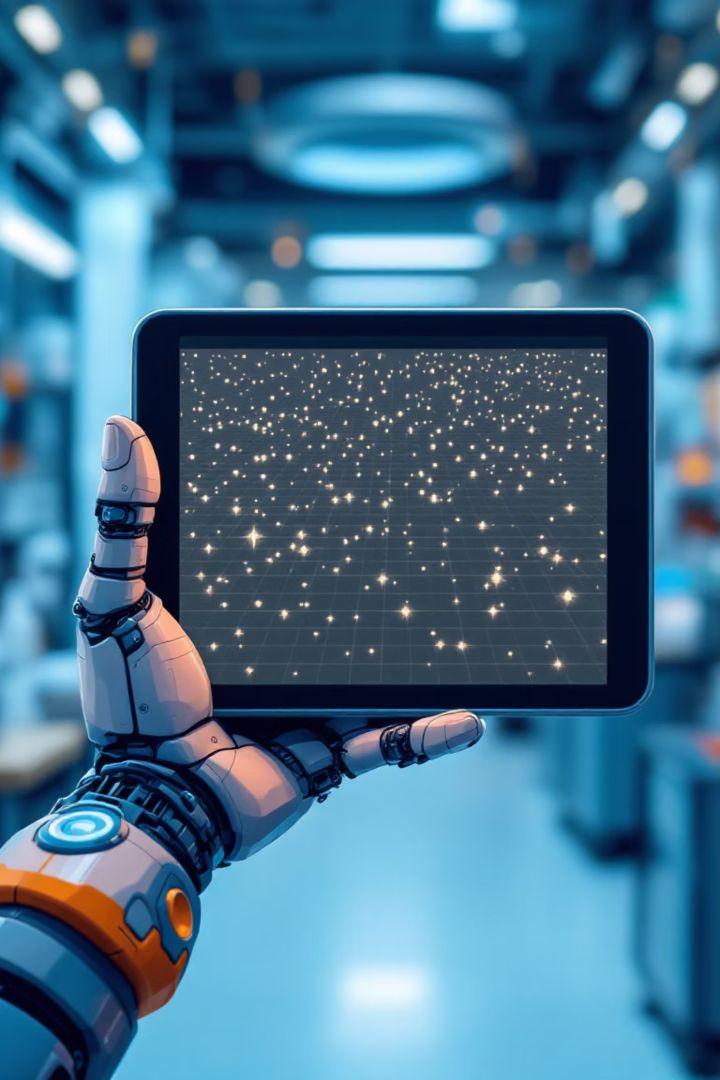
Optimisation TensorRT

Génération d'un moteur d'inférence optimisé pour les plateformes NVIDIA Jetson, utilisant des calculs en précision réduite (FP16/INT8).



Avantages Clés

- Accélération significative de l'inférence.
- Consommation énergétique maîtrisée.
- Capacité à fonctionner en temps réel sur des dispositifs embarqués.



Conclusion et Perspectives

Cette recherche ouvre de nouvelles voies pour des systèmes de perception 3D plus accessibles et performants dans le domaine de la robotique et des véhicules autonomes.

"Combiner la simplicité d'un LiDAR 2D avec la richesse d'une caméra RGB et la puissance de l'IA permet de démocratiser la perception 3D de haute qualité."

Les futures étapes incluront l'exploration de nouvelles architectures, l'intégration de données temporelles et la validation sur des scénarios environnementaux plus complexes.