

# Perception du Véhicule Extrapolée par Apprentissage Auto-Supervisé

## Florent CHIARONI

Véhicule à conduite déléguée:



Directeur de thèse : Sébastien GLASER  
Encadrant VEDECOM : Mohamed-Cherif RAHAL  
Encadrant ISL : Nicolas HUEBER  
Domaine : Vision et Apprentissage Automatique

### Contexte et Objectifs

Dans le cadre d'une collaboration entre les instituts VEDECOM et ISL, ce travail a pour objectif de montrer l'intérêt d'applications d'apprentissage profond de perception multimodales pour des applications destinées au véhicule à conduit déléguée.

L'équipe VEH08 de VEDECOM est spécialisée dans la délégation de conduite de véhicule et le laboratoire ELSI de l'ISL dans les systèmes de perception intelligents et peu coûteux en énergie.

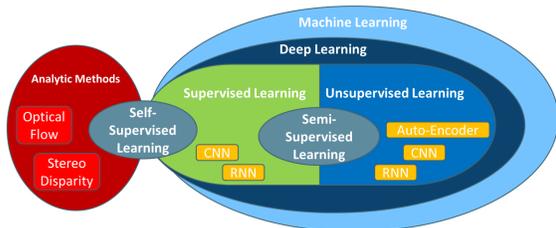
Notre objectif est de détecter les voies traversables (routes, chemins), les obstacles statiques (barrières, bâtiments, arbres, ...) et les obstacles potentiellement mobiles (voitures, cyclistes, piétons, ...) en tenant compte des contraintes suivantes:

- Précision de prédiction ; afin de faire une prise de décision sécurisée au possible dans le contexte du véhicule autonome
- Coût en calcul; pour intégrer la solution dans un système embarqué de basse consommation
- Entraînement réalisable sans données labélisées à la main ; afin de permettre au système la capacité à s'adapter à un éventuel environnement inconnu

### Travaux Antérieurs

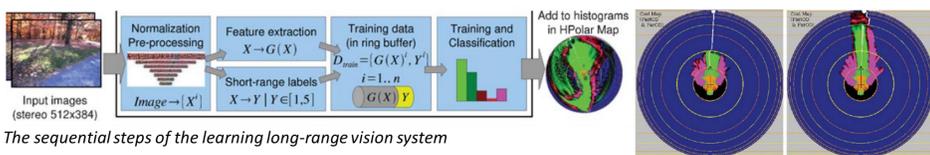
De précédentes étude ont montré l'intérêt de l'apprentissage profond dans le contexte du véhicule autonome afin d'obtenir une meilleure analyse de l'environnement sans le besoin de données labélisées manuellement.

Notions d'apprentissage profond exploitées et leurs interactions :



Articles inspirants pour notre approche :

- **Learning Long-Range Vision for Autonomous Off-Road Driving [1]**. Cet article propose une méthode pour naviguer de manière fluide et rapide dans un environnement tout-terrain inconnu en utilisant une vision à longue portée générée avec une approche d'entraînement auto-supervisé en ligne basé sur des techniques de stéréo vision.



The sequential steps of the learning long-range vision system



Left: Analytic stereovision method output. Right: Learning model output.

Left: Hyperbolic polar map from the analytic stereovision method output. Right: Hyperbolic polar map from the learning model output.

- **Deep tracking : Seeing Beyond Seeing Using Recurrent Neural Networks [2]**. Cette méthode d'apprentissage permet de traquer les obstacles mobiles, même en cas d'occlusion. Le réseau récurrent est ici spécialisé dans le traitement de données séquentielles temporelles en utilisant un état caché dans sa structure.

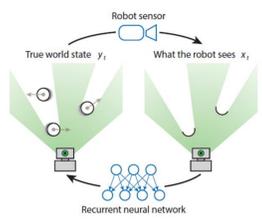
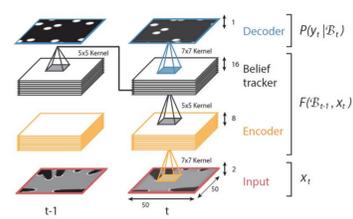
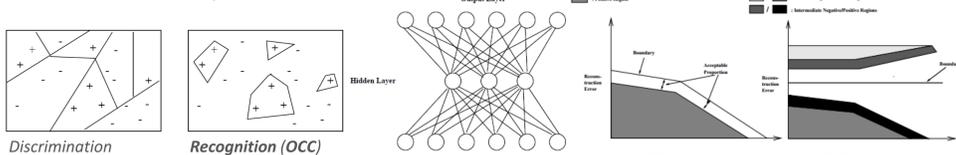


Illustration of the final goal of the method [2]; track moving obstacles, even if they become occluded



Encoder-Recurrent-Decoder structure suggested by [2]

- **Concept-learning in the absence of counter-examples [3]**. Proposition d'un réseau de neurones adapté à l'OCC (One-Class classification) ; l'auto-encodeur. Le modèle est entraîné à reconnaître la classe positive.



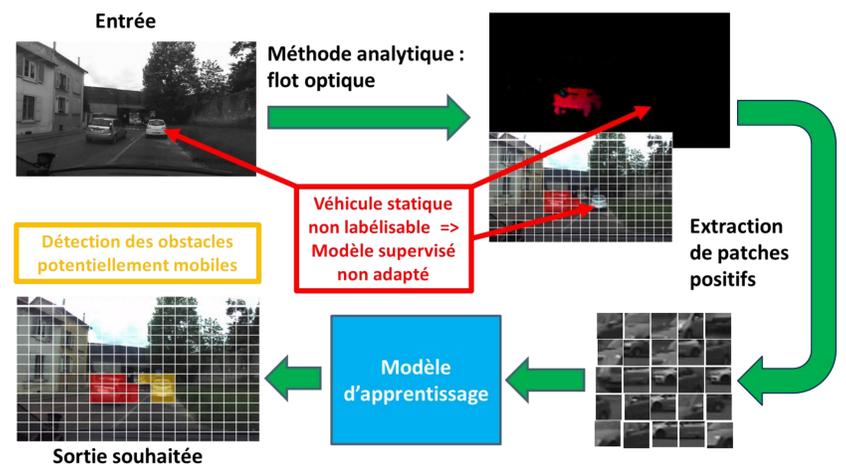
Auto-Encoder structure trained to reconstruct in output its inputs which are only positive samples. Top Boundary of the positive class reconstruction error. The model is trained to reconstruct the positive samples better than the negatives.

- **Learning both weights and connections for efficient neural network [4]**. Cette méthode permet de réduire le coût en calcul d'un réseau de neurones profond ; Le réseau est élagué en apprenant et conservant uniquement ses poids et connexions pertinents.

### Approche proposée

**Apprentissage auto-supervisé pour la détection d'obstacles potentiellement mobiles :**

Ci-dessous l'illustration de notre système auto-supervisé, qui utilisera un modèle d'apprentissage conçu pour détecter les obstacles potentiellement mobiles non détectables par méthodes analytiques traditionnelles telles que le flot optique.

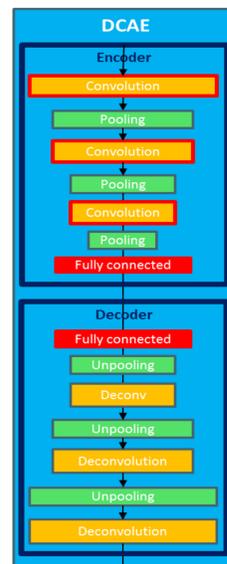


Les données de sortie pourront fournir une analyse de scène plus complète, sans nécessiter de données labélisées manuellement.

**Notre modèle d'apprentissage : DCAE (Denoising Convolutional Auto-Encoder)**

On adapte ici la méthode [3] à l'auto-encodeur convolutif pour permettre la classification en une classe (OCC) sur images.

Résultats du DCAE en mode OCC :



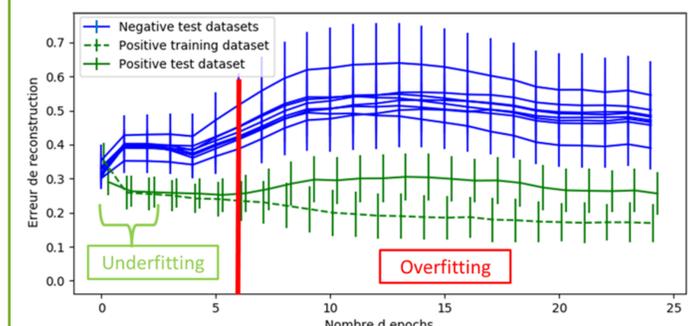
Structure de notre modèle d'apprentissage

Entrée :

Sortie :

- Entraîné sur "0" :
- Entraîné sur "1" :
- Entraîné sur "4" :

DCAE entraîné en mode OCC respectivement sur les classes "0", "1" et "4" du dataset MNIST. On observe que les échantillons de la classe utilisée lors de l'entraînement sont les mieux reconstruits, donc reconnus.



Moyenne et écart-type de l'erreur de reconstruction du DCAE en fonction du nombre d'epochs. Le DCAE est entraîné ici sur la classe positive "0" du dataset MNIST. On observe que l'erreur de reconstruction des échantillons positifs est inférieure à celle des échantillons négatifs.

L'OCC sur images semble possible avec un DCAE. La prochaine étape est de tester cette méthode sur notre système auto-supervisé pour la détection d'obstacles potentiellement mobiles.

### Références

- [1] R. Hadsell et al., "Learning long-range vision for autonomous off-road driving," *Journal of Field Robotics*, vol. 26, no. 2, pp. 120–144, 2009.
- [2] P. Ondruska and I. Posner, "Deep tracking: Seeing beyond seeing using recurrent neural networks," *arXiv preprint arXiv:1602.00991*, 2016.
- [3] N. Japkowicz, *Concept-learning in the absence of counter-examples: an autoassociation-based approach to classification*. Rutgers University, 1999.
- [4] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in *Advances in Neural Information Processing Systems*, 2015, pp. 1135–1143.