

Home Issue & hypothesis Environement Project Conclusion

Introduction to

SELF-DRIVING

WITH CARLA





Agent de conduite autonome basé sur l'apprentissage par renforcement

- Déplacer en autonomie
- Adapter sa vitesse à la densité du traffic
- Gérer l'environnement Carla

[Home](#)[Issue & hypothesis](#)[Environement](#)[Project](#)[Conclusion](#)

- L'utilisation des caméras RGB va être trop gourmande
- Environnement Carla est trop lourd pour effectuer beaucoup d'épisode d'entraînement
- Focaliser sur la qualité de la reward et la sélection du modèle pour apprendre avec le moins d'épisodes possible

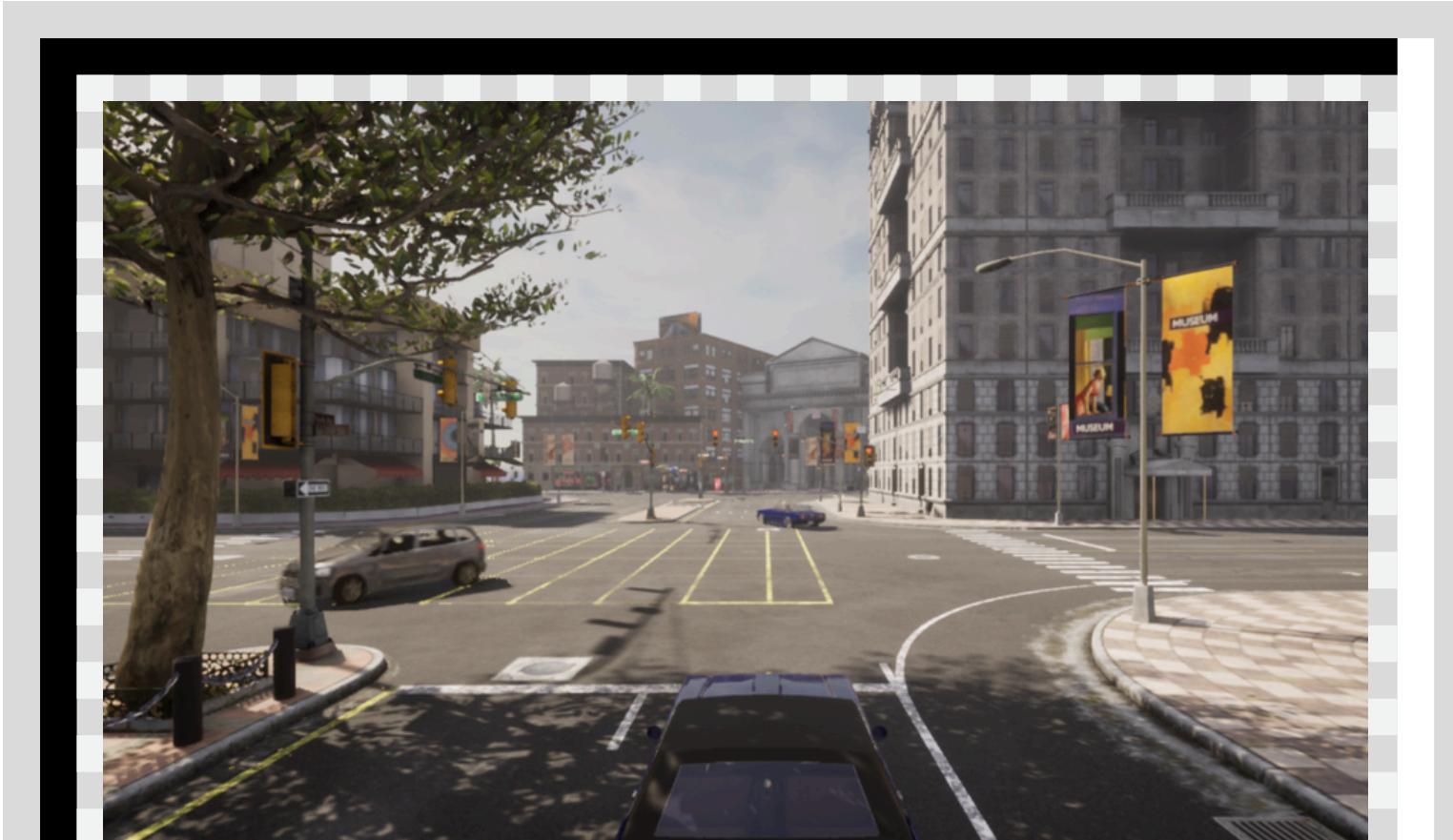




Environnement de conduite urbain virtuel

Utilisation au cours du projet de plusieurs outils nous donnant des observations

- capteur position
- capteurs vitesse
- capteurs collision
- lidar
- waypoint
- caméra rgb



**37 observations finales
(inputs finales)**

- 16 distance mini (lidar)
- 16 distance moyenne (lidar)
- 1 collision
- 1 position
- 1 vitesse
- 1 lane offset
- 1 lane angle



The screenshot shows the CARLA environment's user interface. At the top left is a button labeled "i" (information) with a magnifying glass icon. Next to it is a large button labeled "LES ACTIONS" (Actions). A cursor arrow is pointing at the "LES ACTIONS" button. Below this, there are four blue square buttons arranged in two rows: a top row with a single upward-pointing arrow button, and a bottom row with three buttons: left, down, and right arrows respectively.

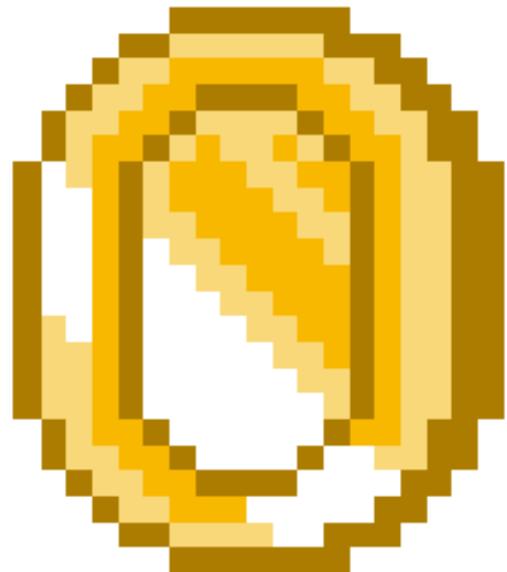
- direction (-1 à 1)
- throttle (0 à 1)
- brake (0 à 1)

At the bottom of the slide, there is a navigation bar with three icons: a left arrow, a right arrow, and a double right arrow.



Récompense à chaque step :

- **Base** : favorise la survie
- **Lane_keeping** : favorise min lane offset et lane_angle
- **Consistency_bonus** : favorise lane Keeping soutenu
- **Exploration** : favorise la distance par rapport au spawn



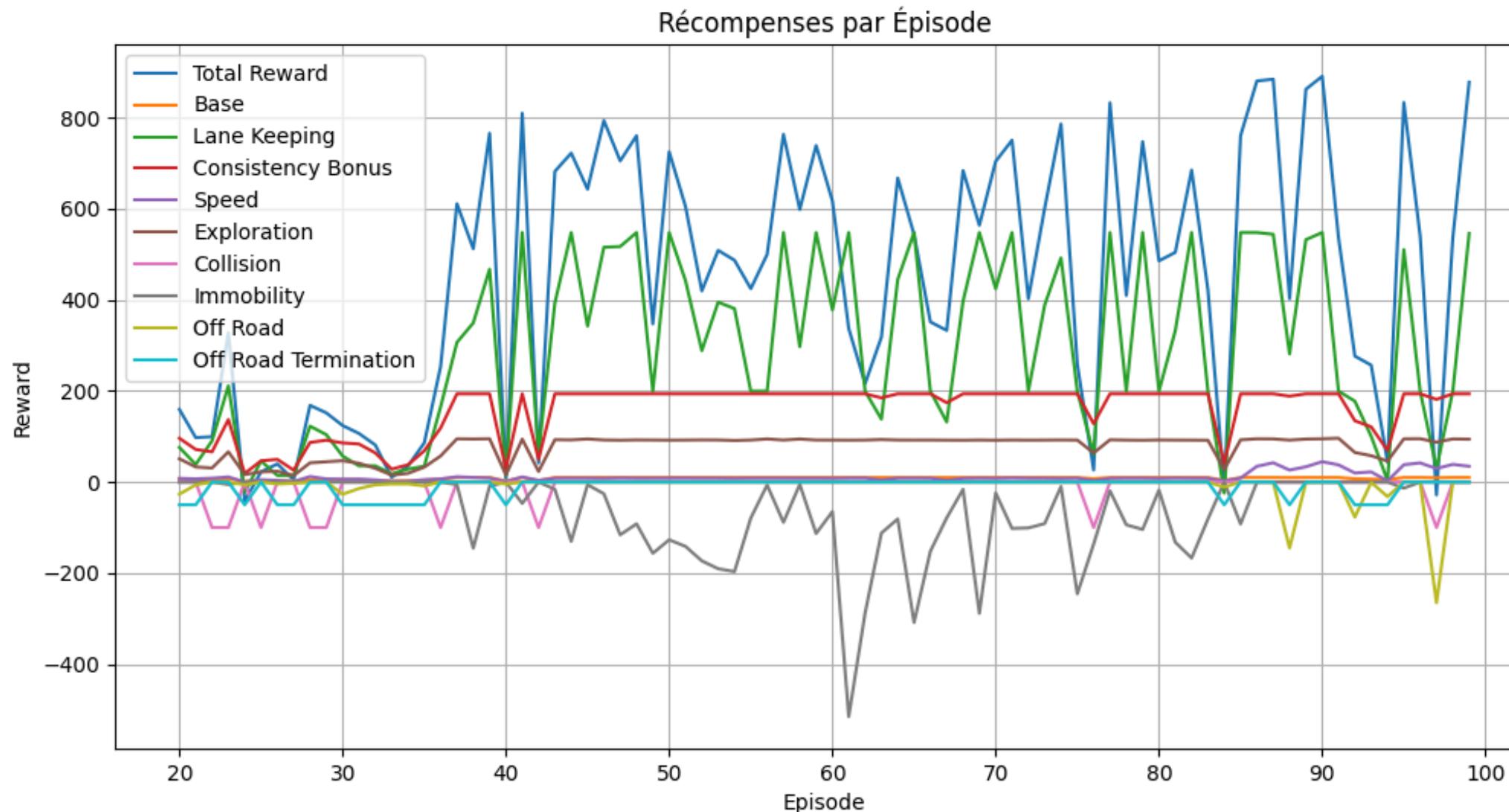


Pénalités critiques :

- **Collision** : pénalise les collisions
- **Immobility** : pénalise l'immobilité excessive
- **Off_road** : pénalise nb step trop loin ou desaxé de la voie
- **off_road_terminaison** : pénalise fortement sortie de route



EVOLUTION RÉCOMPENSE

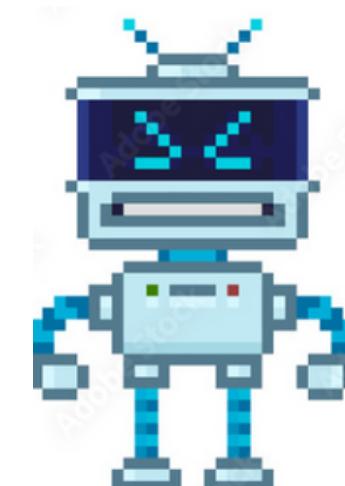
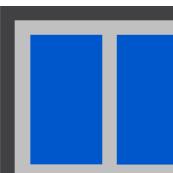


Mise en place d'un clipping pour stabiliser la récompense à chaque step et éviter les valeurs extrême, non représenté dans le graphique

Le graph montre entraînement d'un modèle sélectionné pour sa convergence rapide après les 10 premiers épisodes



EVOLUTION DU PROJET & MODÈLE

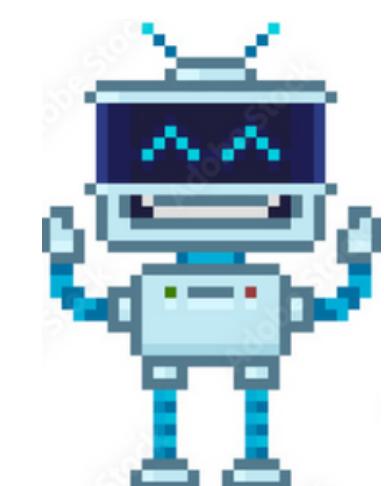


DÉCOUVERTE DE CARLA

- Etude de l'environnement
- Contrôle de Carla et Blueprint
- Spawn voiture, caméra et contrôle

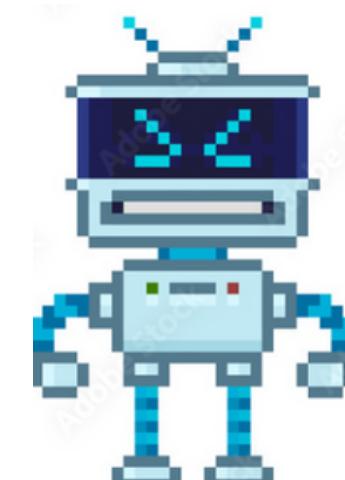
Difficulté à faire tournée
l'environnement sur nos machines

Entrainement pas facile mais optimisable



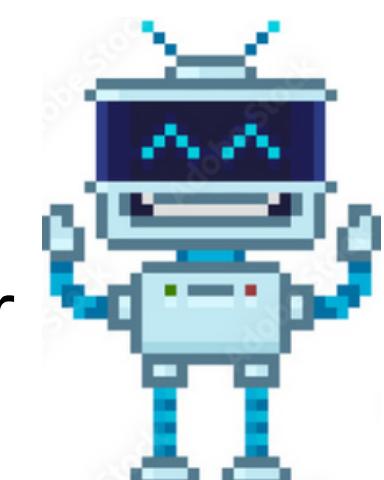


EVOLUTION DU PROJET & MODÈLE



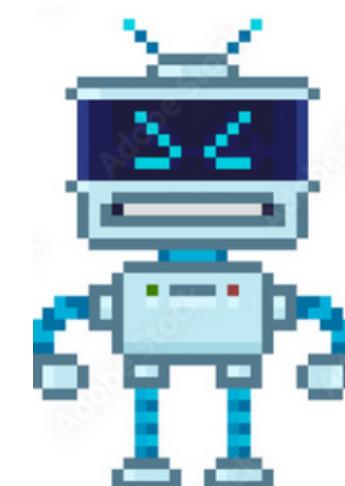
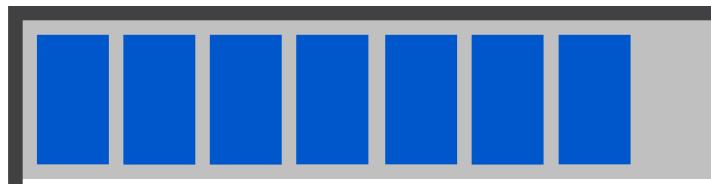
Difficultée de trouver de la performance avec DQN
Modèle vue en cours insuffisant dans notre situation pour la V0

Nécessaire de chercher un modèle plus efficient car entraînement DQN insuportable par nos machines



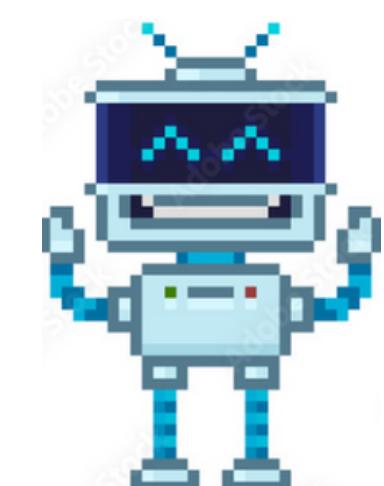


EVOLUTION DU PROJET & MODÈLE



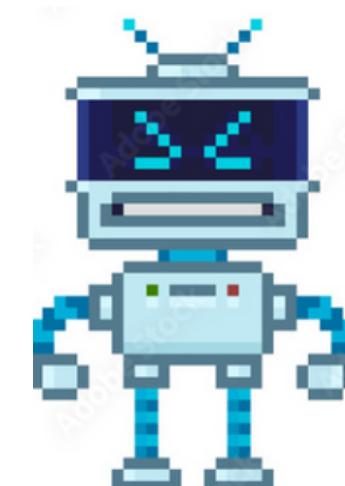
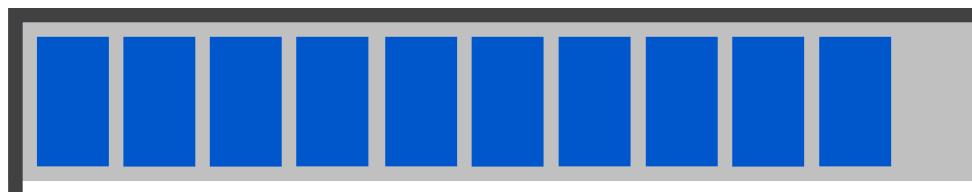
Iteration d'entraînement pour rechercher l'efficience mais toujours pas mal de difficulté. Perte de motivation car loin du livrable finale et la convergence de nos modèles résulte plus de la chance que de l'optimisation de la reward

Passage de la v0





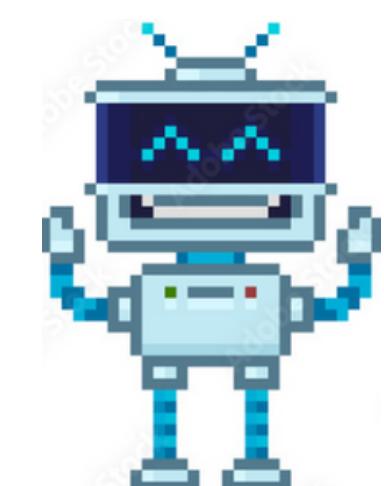
EVOLUTION DU PROJET & MODÈLE

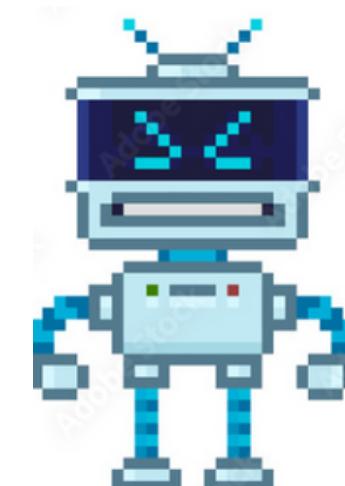
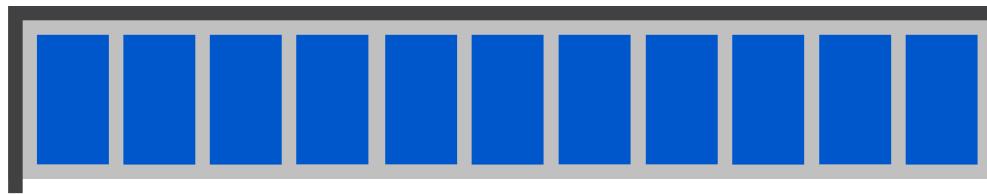


DERNIÈRE OPTIMISATION

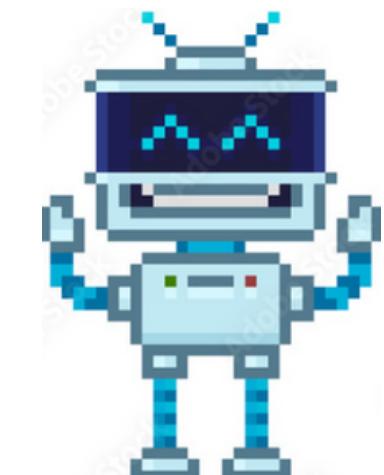
- Modification légère des poids de la reward
- Lancement d'entraînement en boucle pour espérer avoir le modèle le plus performant possible
- Essai de la V1

Plus le modèle est performant plus sa vitesse et les collisions qu'il prend sont intenses en s'entraînant
=> Arrête l'entraînement car le pc bloque à cause des calculs





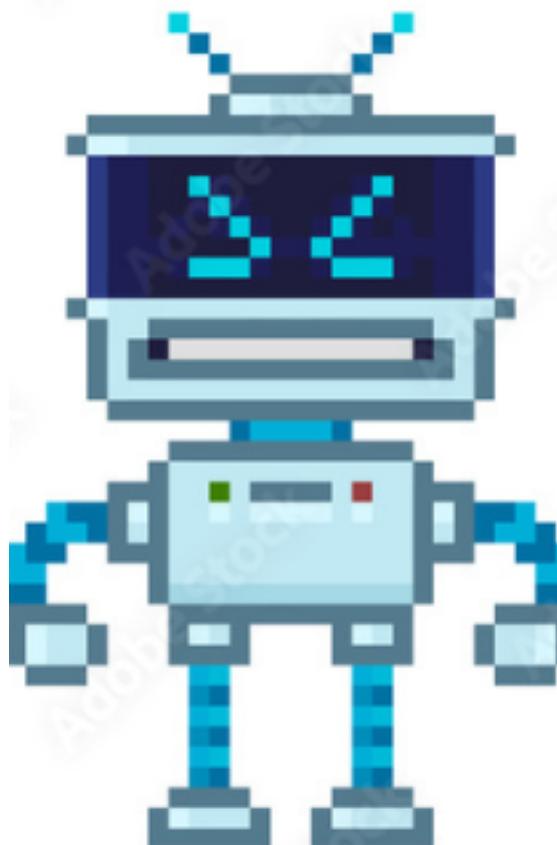
- Modèle de **policy gradient** plus pertinent = output en valeur continu (pas discréétisation = +précis et +exploration)
 - Modèle de type **on-policy** (réduit la divergence)
 - Très stable avec du **clipping**
 - Input only (non compatible avec image)
 - Nécessite bien moins d'input et d'épisode pour converger
 - Plus facile à implémenter (SAC / TD3)
- = **Moins couteux en ressource**



Home Issue & hypothesis Environement Project Conclusion



RÉSULTATS

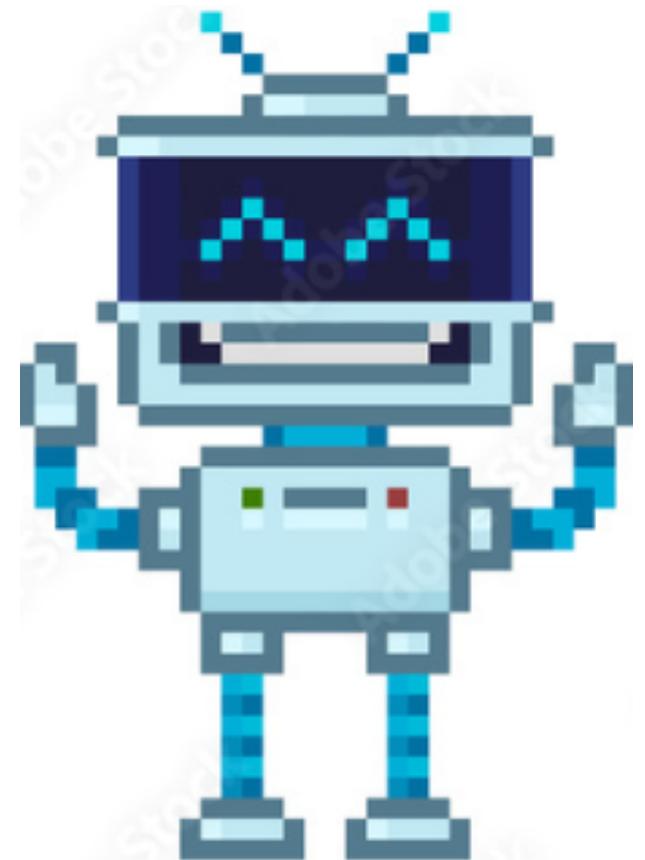


Notre agent fait:

- Suivre la route
- Ne change pas de voie
- des vecteurs d'observations, un ensemble d'actions, une fonction de récompense

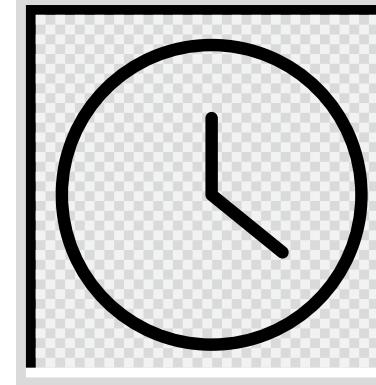
Ne fait pas :

- Suivre les virages
- Correctement adapter sa vitesse face aux obstacles
- changer de voie

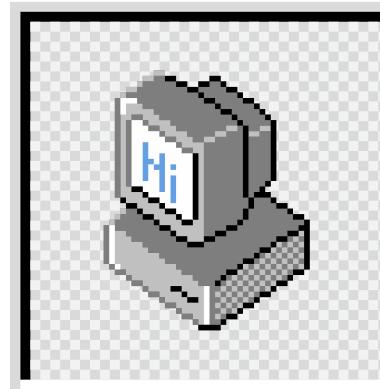




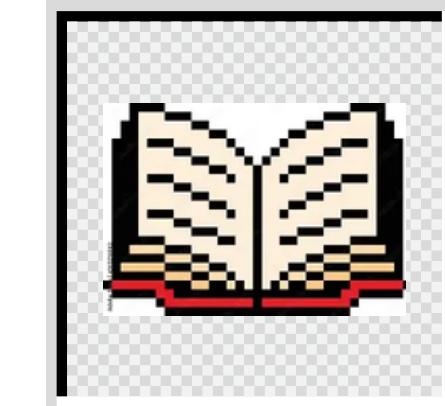
Manque de temps



Manque de CPU & GPU



Manque de compétence



CONCLUSION

Projet intéressant mais trop prétentieux pour les ressources disponible

Les résultats sont objectivement pas suffisant

Travail honorable au vue des ressources allouée

Louis LAMBERT

Théo CONDAMIN

