

Finalisation du Design de l'Amélioration

Catégorie	Description
Tâches réalisées et finalisées	<ul style="list-style-type: none">- Conception du Design Final de l'Agent RL : Espace d'état défini comme un vecteur d'opérations. Espace d'action défini comme la modification/sélection d'une opération- Fonction de Récompense : Conçue pour maximiser l'amélioration du rang prédit par les proxys zéro-coût. Une architecture avec un meilleur classement que la précédente reçoit une récompense positive.- Préparation du PPO : Identification et adaptation du code pour un agent PPO simple, qui sera pré-entraîné sur une tâche proxy.
Tâches en cours de réalisation	<ul style="list-style-type: none">- Finalisation de l'architecture du module RL et tests unitaires pour assurer que les entrées/sorties correspondent aux interfaces de RD-NAS.
Tâches qui vont démarrer la semaine prochaine	<ul style="list-style-type: none">- Démarrage de l'implémentation : Intégration de l'agent PPO pour remplacer la méthode d'échantillonnage aléatoire d'architectures dans la boucle de recherche de RD-NAS.
Risques et difficultés	<ul style="list-style-type: none">- Instabilité de la Récompense RL (Niveau Modéré) : La fonction de récompense, bien que basée sur une métrique forte (le rang), peut être bruitée. Cela pourrait rendre l'apprentissage de l'agent RL instable et nécessiter un ajustement fin des hyperparamètres du PPO.
Lien GitHub	https://github.com/waelsouissi2/RD-NAS_enhancement