

Implémentation de l'Agent de Recherche (RL)

Catégorie	Description
Tâches réalisées et finalisées	<ul style="list-style-type: none">- Implémentation du PPO simplifié : Intégration de la structure de l'agent Proximal Policy Optimization (PPO) (inspirée du Papier 2, mais en version classique dans un premier temps) pour gérer la politique de recherche d'architecture.- Remplacement de la logique de sélection : Modification du module d'échantillonnage de RD-NAS pour utiliser les actions générées par l'agent RL au lieu de l'échantillonnage aléatoire.- Débogage des communications entre l'agent RL (qui choisit l'architecture) et la logique de calcul du proxy (qui fournit la récompense).
Tâches en cours de réalisation	<ul style="list-style-type: none">- Test d'un premier cycle complet de l'agent PPO : sélection d'une architecture, calcul du score de proxy, calcul de la récompense, et mise à jour de la politique.
Tâches qui vont démarrer la semaine prochaine	<ul style="list-style-type: none">- Mise en œuvre du Transfer Learning : Préparation du jeu de données proxy (e.g., une tâche simple comme CIFAR-100 ou un sous-ensemble) pour pré-entraîner l'agent RL.
Risques et difficultés	<ul style="list-style-type: none">- Problèmes de Convergence (Niveau Modéré) : L'algorithme PPO est sensible aux hyperparamètres. Le processus d'intégration pourrait introduire de l'instabilité, nécessitant un ajustement précis des paramètres d'apprentissage (taux d'apprentissage, K epochs, clipping).
Lien GitHub	https://github.com/waelsouissi2/RD-NAS_enhancement