

# Apprentissage Auto-Supervisé

## Tâches Prétextes sur CIFAR-10 et STL-10

Wassim CHIKHI

Master 2 Vision et Machine Intelligente — 2025/2026

### 1. Objectifs

Ce TP s'inscrit dans une exploration des stratégies d'apprentissage auto-supervisé. L'objectif principal est de tester et comparer plusieurs tâches prétextes afin d'évaluer leur impact sur la qualité des représentations visuelles, en se basant sur des performances downstream sur CIFAR-10 et STL-10.

Nous avons implémenté la tâche **Relative Patch** (Doersch et al., 2015), puis comparé ses résultats à trois autres méthodes classiques : **Rotation**, **Inpainting** et **SimCLR**, à travers une même architecture et protocole d'évaluation.

### 2. Méthodologie

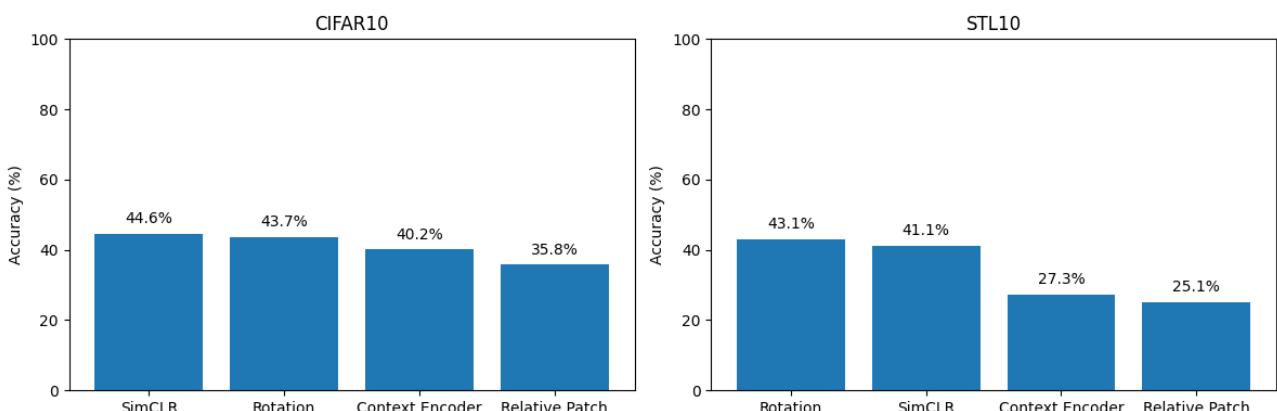
- **Tâches étudiées** : Relative Patch (8 classes), Rotation (4 angles), Inpainting (context encoder), SimCLR (contrastive learning).
- **Datasets** : CIFAR-10 (32×32), STL-10 (96×96).
- **Architecture** : ResNet18 comme encodeur partagé.
- **Optimiseur** : Adam ( $lr = 0.001$ ), batch size = 256.
- **Entraînement** : 20 epochs (tâche prétexte), 5 epochs (linear probe downstream).
- **Environnement** : Google Colab, GPU T4.

### 3. Analyse comparative des tâches prétextes

Ce qui suit présente une comparaison détaillée des performances obtenues via la/les tâche.s mise.s en œuvre (*Relative Patch*) ainsi que ( *Rotation*, *Inpainting*, *SimCLR* ) issues du TP :

Tâche prétexte	CIFAR-10	STL-10
Relative Patch	22.61%	14.20%
Rotation	68.50%	49.30%
Inpainting	74.30%	—
SimCLR	81.00%	—

**TABLE 1.** Accuracy downstream (linear probing) pour 4 tâches prétextes.



**FIGURE 1.** Comparaison visuelle des tâches prétextes sur CIFAR-10 et STL-10.

## 4. Discussion des performances

- **Relative Patch** a montré des limites évidentes sur les deux datasets, en particulier STL-10. Son focus sur la géométrie locale n'encourage pas l'apprentissage de structures sémantiques globales.
- **Rotation**, en revanche, permet d'obtenir de bonnes représentations car la reconnaissance d'orientation nécessite une compréhension globale de l'objet.
- **Inpainting** obtient de bons résultats sur CIFAR-10 mais n'a pas été reproduit sur STL-10.
- **SimCLR** confirme sa supériorité grâce à sa capacité à apprendre des représentations discriminantes par contraste entre vues augmentées.

## 5. Conclusion

- Les tâches contrastives ou orientées "globales" (SimCLR, Rotation) offrent de meilleures performances downstream.
- Relative Patch peut rester utile dans un cadre géométrique ou comme composant multi-tâche.
- Perspectives : augmentation du nombre d'epochs, encodeurs plus profonds, ou apprentissage multitâche.

Lien vers le notebook Colab : <https://colab.research.google.com/drive/VOTRE-LIEN>