

Benchmarking Canonical Evolution Strategies for Playing Atari

Auteurs : Patryk Chrabaszcz, Ilya Loshchilov, Frank Hutter

Présenté par : Denis Zheng / Petar Calic

Contexte de l'article

- Volodymyr Mnih (2013) : Deep RL capable de gérer des problèmes avec entrées de grandes dimensions (DQN)
- Tim Salimans (2017) : Les stratégies d'évolution naturelle (NES) peuvent concurrencer les états de l'art de RL
- Patryck Chrabaszcz (2018) : Canonical ES peut également concurrencer les stratégies d'évolution naturelle pour Atari

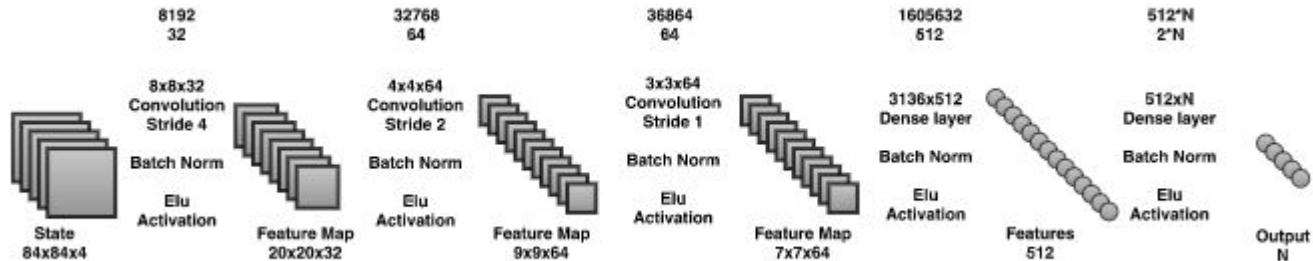
Configuration et résultats

- Pipeline de preprocessing de Mnih :



Configuration et résultats

- Réseau de neurones :



Configuration et résultats

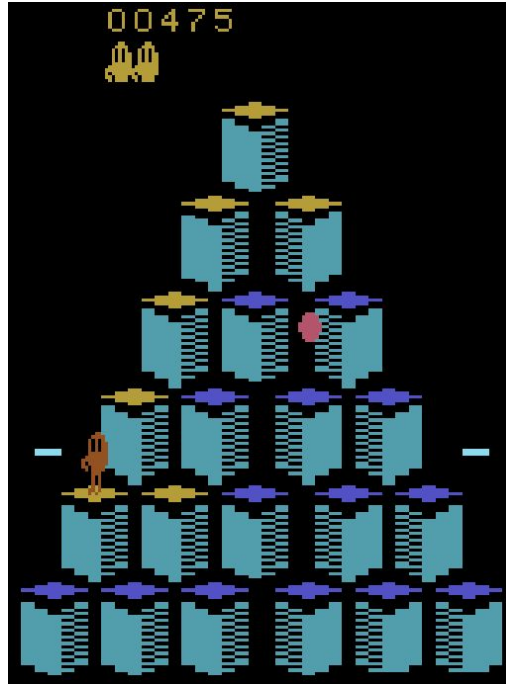
- Training :
 - pour chaque jeu et chaque algorithme, 3 runs
 - chacun sur 400 CPU et avec un budget de 10 heures
 - taille de la population parent fixée à 50
 - chaque run évaluée 30 fois
- Résultats :

Après 5h, sur 24 runs, CES meilleur que OES sur 9, pire sur 7 et 8 égalités

Les auteurs remarquent une nette augmentation des performances entre 1h et 5h de training

Nos résultats

- Jeu choisi : Qbert



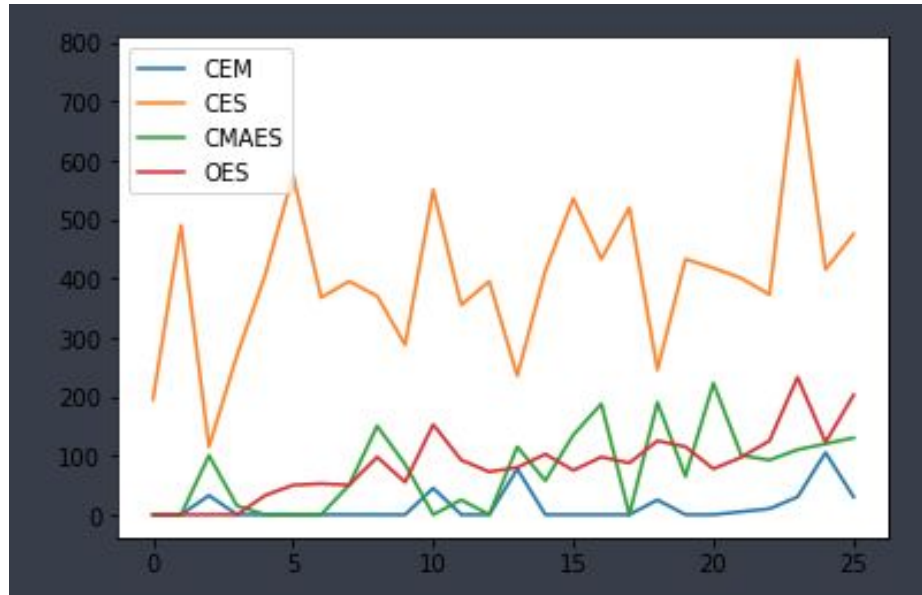
Nos résultats

- Plus petit budget :
 - VM de 8 CPU et 32 GB de mémoire
 - Taille de population des offspring fixée à 16 au lieu de 798
 - Taille de population parent fixée à 2 au lieu de 50
 - 10 évaluations au lieu de 30
- 4 algorithmes : CES, OES, CMA-ES, CEM
- 1 training de 2h et 1 training de 10h par algorithme

Nos résultats

2h

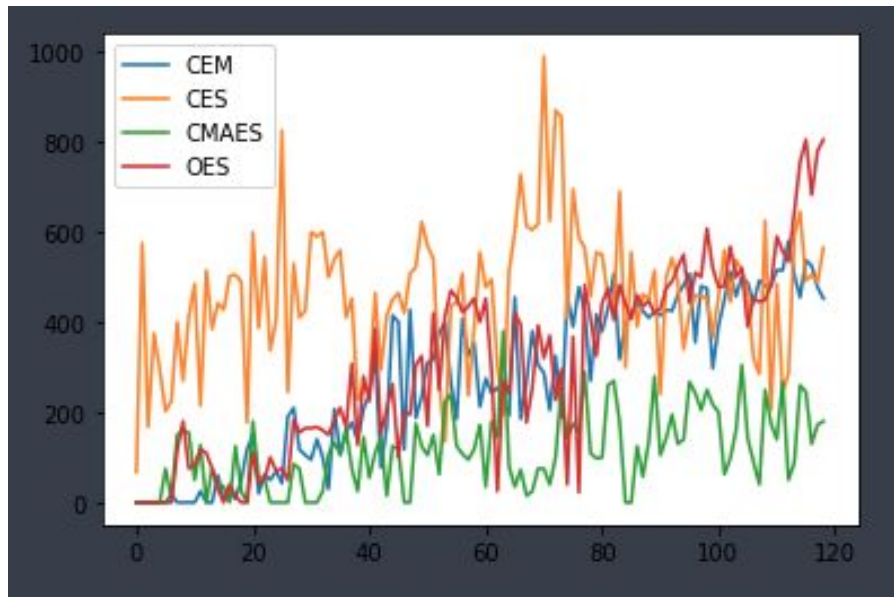
Plot de la fitness moyenne de la population à chaque itération



Nos résultats

10h

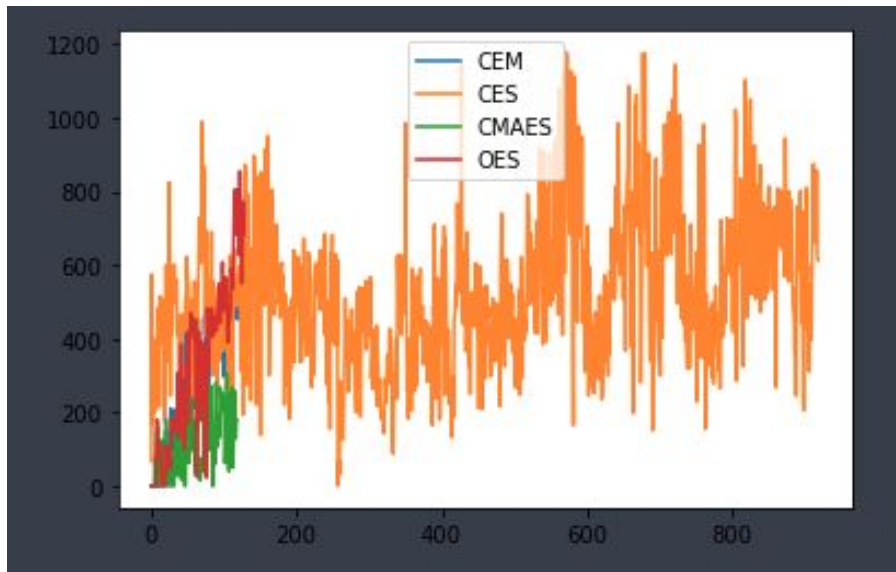
Plot de la fitness moyenne de la population à chaque itération



Nos résultats

10h

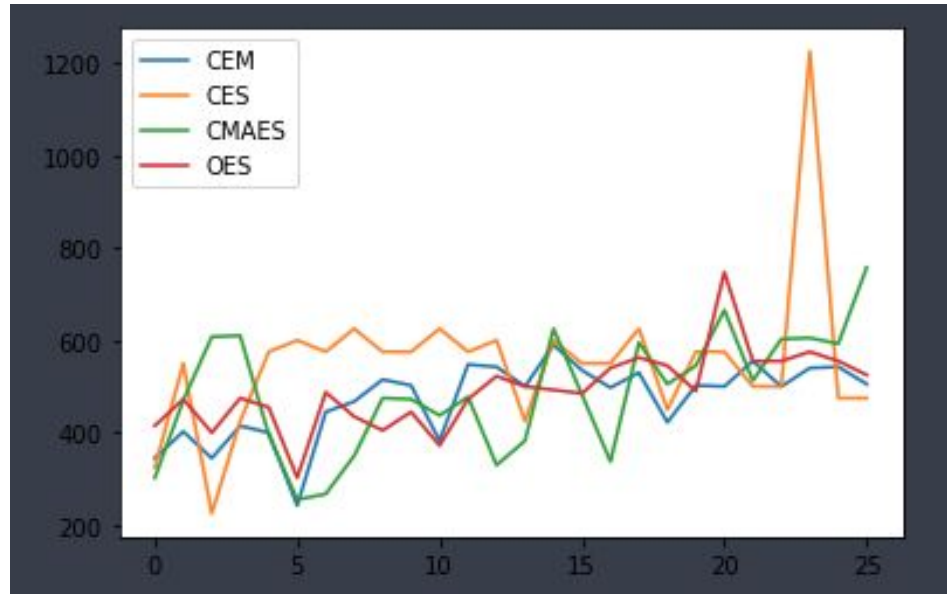
Plot de la fitness moyenne de la population à chaque itération



Nos résultats

2h

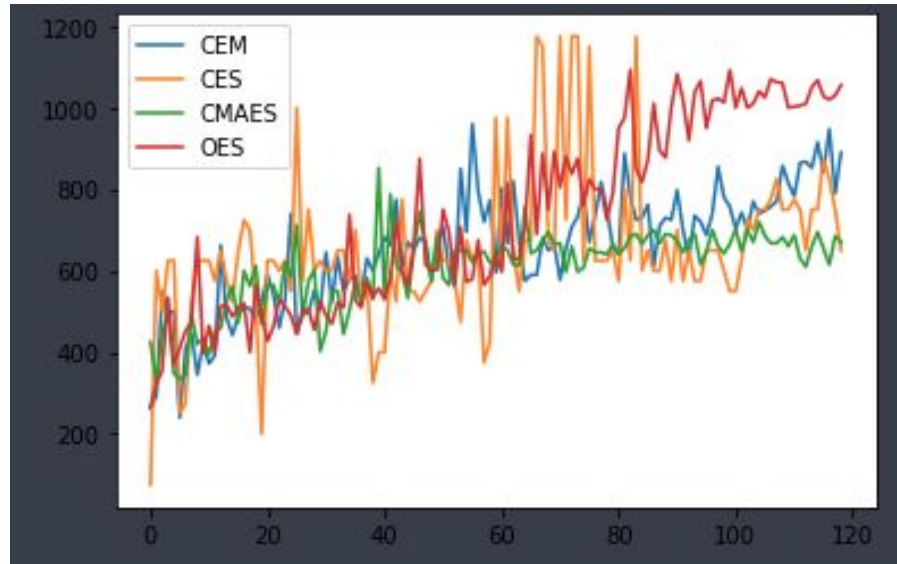
Plot de la fitness du meilleur individu à chaque itération



Nos résultats

10h

Plot de la fitness du meilleur individu à chaque itération



Nos résultats

Résultats de l'article

	OpenAI (2h)	CES (2h)	OpenAI (10h)	CES (10h)
Qbert	8275 1400 1250	8000 6625 5850	12775 5075 4300	263242 16673.3 5136.7

Nos résultats

	OpenAI (2h)	CES (2h)	OpenAI (10h)	CES (10h)
Qbert	491.8	552.8	730.1	653.5

Nos résultats

	OpenAI(2h)	CES(2h)	CMAES(2h)	CEM(2h)	OpenAI(10h)	CES(10h)	CMAES(10h)	CEM(10h)
Qbert	491.8	552.8	486.8	472	730.1	653.5	605.8	654.5

Conclusion

- Nos observations :
 - CES est meilleur à 2h de run, OpenAI est meilleur à 10h
 - CMAES est le plus mauvais dans tous les cas
- CES est bien compétitif avec NES
- Résultats sur Qbert pas forcément reproductibles avec un plus petit budget

Améliorations possibles :

- Faire plus de run
- Hyperparameters tuning (par ex: Optuna)
- Tester sur un jeu plus facile