

進捗報告

1 今週やったこと

- GA の改良

2 実験

2.1 問題

後の順番で学習した個体が有利になるという欠点を修正するため, w と α の訓練を分離した. また適応度を正確にするため, 実際にサンプリングした α とその時点の w で損失を計算した. 以下が改良した GA の手順.

1. 一様乱数で初期個体生成
2. 重み w を $\sum_{i \in P} \nabla_w \mathcal{L}_{\text{train}}(w^*, \alpha_i^{\text{sampled}})$ で更新
3. 個体 α を $\nabla_{\alpha} \mathcal{L}_{\text{valid}}(w^*, \alpha_i)$ で更新
4. 適応度 $\mathcal{L}_{\text{test}}(w, \alpha^{\text{sampled}})$ で個体 α を評価・選択
5. 交叉・突然変異
6. 収束するまで 2. に戻る

(P は個体群, α^{sampled} は隣接行列にサンプリングした α)

α の個体表現と交叉 交叉は行列の同じ位置を参照する一様交叉を採用した. 遺伝子が実数なので, ブレンド交叉も試したが優良解が壊れやすいので導入するときはエリート保存戦略と組合せたい.

2.2 実験設定

表 1, 2 にモデルと GA の設定を示した. 初期収束を回避するため, トーナメントサイズや交叉アルゴリズムを変更した.

初期個体は各辺に $[0,1]$ の一様乱数を与えた.

表 1: モデルの設定

base model	VGG19
Optim(w)	SGD(lr=0.001, momentum=0.9)
Optim(α)	Adam(lr=0.003, $\beta=(0.5, 0.999)$)
Loss	Cross Entropy Loss
dataset	cifar10
pretrain	true
batch size	64
train size	12500
valid size	5000

表 2: GA の設定

個体数	10
世代数	20
選択	トーナメント
サイズ	2
交叉	一様交叉
交叉率	0.5
変異	ガウス分布
変異率 (遺伝子座ごと)	0.2 0.1

2.3 結果

図 2, 3 に GA の結果の精度とロスを示した.

図 1 には個体群のショートカット数を示した. ショートカットが多いほうが精度が高くなりそうだが, 8 本程度に収束している. 同時期の精度や損失を見ると大きく向上しているので, 第一段階としてショートカット本数の学習ができたと思われる.

3 今後の予定

来週は発表の資料を作成する.

ssh のエラーに実験を中断されたが, 次回はこれを修正してさらに大規模に実験する.

Listing 1: error log

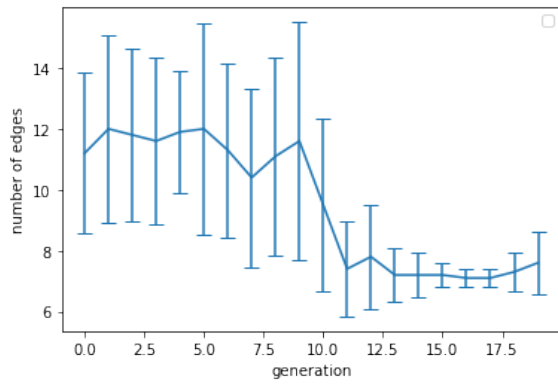


図 1: 世代ごとのショートカット数

```
1 client_loop: send disconnect: Connection
  reset by peer
```

4 ソースコード

github の notebook リポジトリ参照.

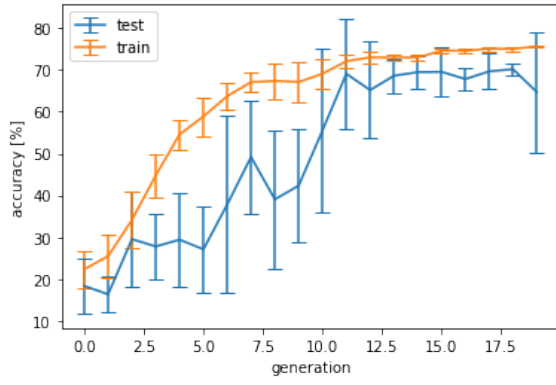


図 2: 世代ごとの精度 (平均と標準偏差)

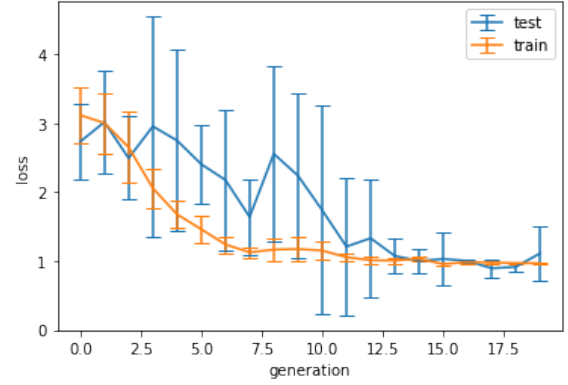


図 3: 世代ごとのロス (平均と標準偏差)

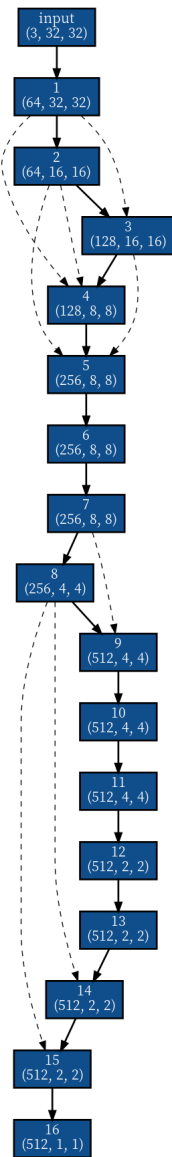


図 4: 1 世代目の最良個体

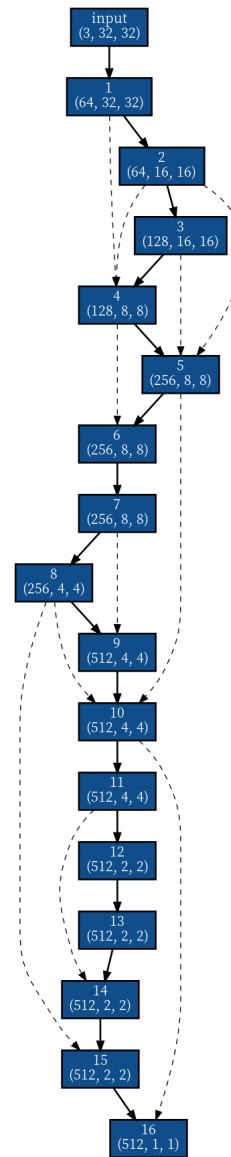


図 5: 20 世代目の最良個体