

進捗報告

表 1: 実験の設定

Cell Depth	5
Node	7(input=2, output=1)
Optim(model)	SGD(lr=2.0e-2, momentum=0.9)
Optim(θ)	Adam(lr=2e-4, $\beta=(0.5, 0.999)$)
Loss	Cross Entropy Loss
batch size	64
train data	25000
epoch	25

表 2: テスト精度の結果

	epoch	Accuracy(%)
アーキテクチャ探索	25	83.6
モデル再学習	25	83.4

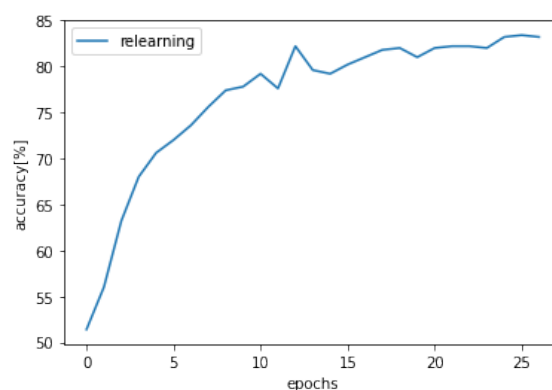


図 1: 再学習時の精度

1 今週やったこと

- コードの修正

2 前回までの問題点

パラメータを持たない Average Pooling が偏って選ばれ、探索したネットワークの精度が出なかった。

3 コードの見直し

コードを確認して以下の修正を行った。

- データセットがシャッフルされていなかったのを、修正
- レイヤーの重みとネットワークの重みを独立したデータセットでそれぞれ学習

4 実験

訓練データはレイヤーの重みとネットワークの重みそれぞれに対して 25000 枚とした。学習時間は 50epoch としたが、自分で設定していた最大学習時間の 3 時間に引っかかってしまったので 25 epoch までを結果として載せる。

5 結果

表 1 には両学習のテスト精度の結果を、図 1 にはモデルの再学習時の精度の経過を示した。図 2, 3 には探索の結果得られたセルの構造を示した。

セルは Convolution と Pooling が共存したバランスとなり、偏りは解消されている。それに伴って、精度も 10% 程度向上した。再学習をもう少し長くすればさらに伸びる可能性がある。元論文では 24 時間学習で精度 90% なので、ある程度近づくことができたといえる。

6 考察

アーキテクチャの更新に unrolled の評価関数を使用していた。unrolled GAN で提案された手法。モデルを更新して先読みしたパラメータで評価するらしいが、ランタイムエラーで実行できなかった (メモリ不足?)。

またメモリ不足の問題だが、モデル自体は 1MB 程度思ったより圧迫していないので、高い batch size で画像データを読み込んでいることが原因だと思われるが、はっきりとした理由や解決法はまだ分かっていない。

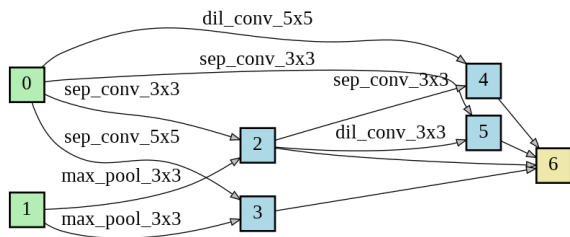


図 2: 学習したセル : normal

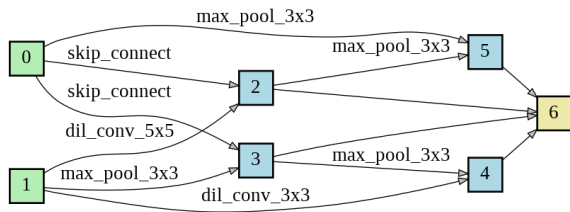


図 3: 学習したセル : reduce

7 今後の予定

- 既存の分類モデルの調査とそれを探索空間に含むアーキテクチャサーチの設計
- (unrolled GAN の論文を読む?)

8 ソースコード

Github の同階層の `NAS_test.ipynb` を参照してください.