

進捗報告

表 1: 実験の設定

Cell	6
Node	7(input=2, output=1)
Optim(model)	SGD(lr=2.5e-2, momentum=0.9)
Optim(θ)	Adam(lr=2e-4, β =(0.5, 0.999))
Loss	Cross Entropy Loss
batch size	64
train data	20000
epoch	10+40+50

表 2: 演算子候補

none
skip_connect
max_pool_3x3
sep_conv_3x3
sep_conv_5x5
dil_conv_3x3
dil_conv_5x5

1 今週やったこと

- セルの多入力対応
- 演算子追加

2 NAS

2.1 設定

表 1 には実験設定を示した。入力・出力ノードの数は、ともに 1 に設定した。また出力ノードへの接続はチャンネルの concat であり、今回は 4 つのノードを使ってチャンネル数を 4 倍にした。ノードは 7 にしたため、探索する辺は 14 となった。表 2 のように、畳み込み層、プーリング層、恒等写像、零写像の 6 つの演算子を用意した。またセルの入力は、チャンネル数の前処理として Pointwise Convolution を用いた。

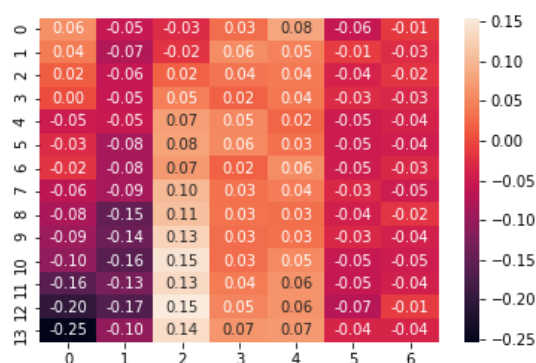
このセルを 4 つ重ねたものを用いて、Cifar10 の 10 クラス分類器を構築した。モデルの Optimizer は SGD で、アーキテクチャを表す。

2.2 変更点

1 入力から多入力に拡張した。2 入力以上でも実験できるようにした。

Separable Conv と Dilation Conv を追加し、演算子の数を増やした。

前処理による初期チャンネル数を 16 に増やした。

図 1: θ 縦軸は辺の番号, 横軸は演算子の番号

2.3 実験

実験ではまず (a)10 epoch 学習し、その後 (b)40 epoch 訓練した。得られた、重みを (c)50 epoch で再学習した。訓練時間は全体でおよそ 3 時間程度であった。

図 1 にはノーマルセルの θ を示した。縦軸は順に (0, 2), (0, 3), (0, 4), (0, 5), (1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5) の辺を、横軸は表 2 の演算子のインデックスを表している。

図 2, 3 に得られたセルを示した。

3 考察

論文通りの演算子を追加して、入力も合わせたが、max pooling ばかりが選ばれ精度は 75% となった。コードを見直してバグがないかや、論文の設定を調査したい。

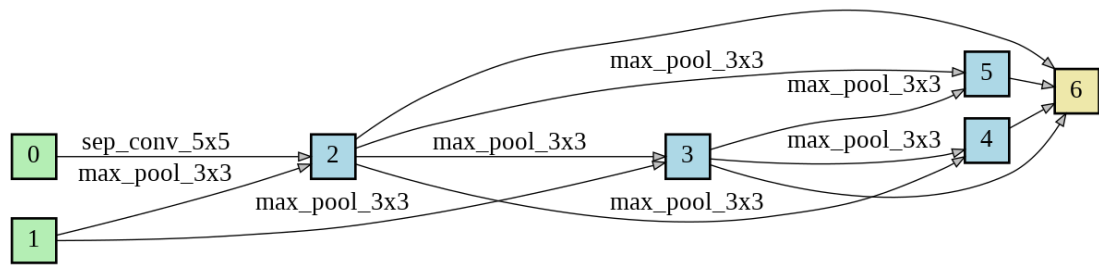


図 2: Normal Cell

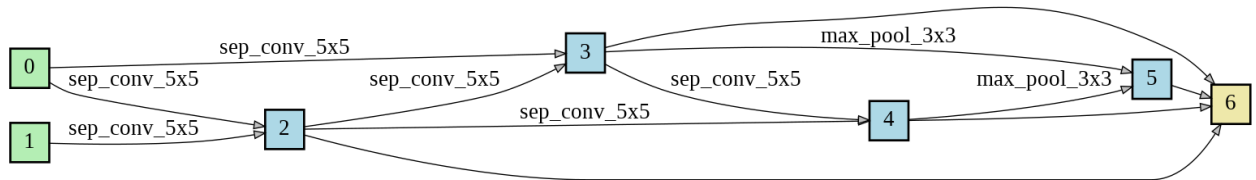


図 3: Reduction Cell

現在使用している Google Colab のホストマシンで RAM が不足する. より大きな設定の実験はローカルで実行するか, メモリの解放ミスを確認する.

4 今後の予定

- コードの確認
- ローカルサーバーでの実行

5 ソースコード

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