

進捗報告

表 1: モデルの設定

base model	VGG19
Optim(w)	SGD(lr=0.001, momentum=0.9)
Optim(α)	Adam(lr=0.001, $\beta=(0.5, 0.999)$)
Loss	Cross Entropy Loss
dataset	cifar10
pretrain	true
batch size	64
train size	25000
valid size	25000

表 2: 結果

タイプ	accuracy(%)	alpha #
A	93.71	17 * 17
B	93.74	17 * 2
C	93.80	17 * 3 - 1

4 今後の予定

- 今回の 3 つの手法で TDGA の実験をする

1 今週やったこと

- α の改良

2 実験

図 1, 2, 3 に従来の α の形式 A とベクトルの積で行列を作る タイプ B, C のイメージを示した.

タイプ B は学習する変数が少なく, タイプ C は縦方向の深さと斜め方向の接続距離の積で, 構造的な特徴を反映できるようにした.

表 1, にモデルの設定を示した.

3 結果

表 2 に 3 つのタイプのネットワーク性能と学習した変数の数を示した. タイプ C は従来のタイプ A と比較して 0.09% 向上した. 最も変数が少ない B よりも C がよりよい結果となった. タイプ C の接続距離を考慮した構造がうまく作用したと思われる.

図 4, 5, 6 に行列にしたときの α の重みを示した. タイプ A ではピンポイントな位置の推定ができていて一方で, タイプ B, C は重みは直線上は似た値になっており, 大雑把な設定であったかもしれない. それにも関わらず良い結果となったのは, 構造推定に関して主観的な拘束条件を加えたからであると考えられる. しかし深さに対して $O(n^2)$ だった問題を解決する方法の 1 つは示せた.

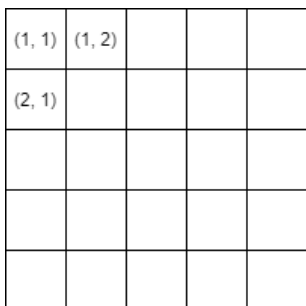


图 1: Type A

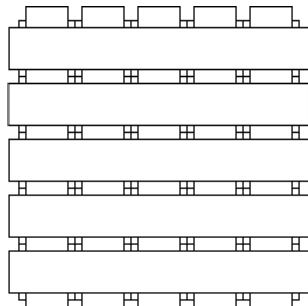


图 2: Type B

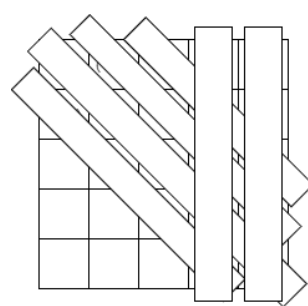


图 3: Type C

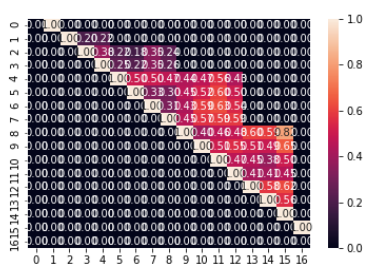


图 4: Type A

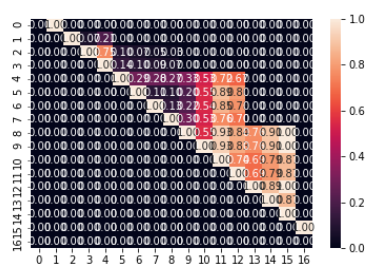


图 5: Type B

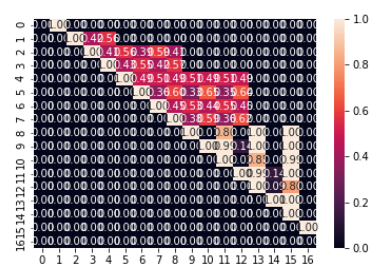


图 6: Type C