

## 進捗報告

表 1: 実験の設定

base model	VGG19
Optim( $w$ )	SGD(lr=0.0090131, momentum=0.9)
Scheduler( $w$ )	Step( $\gamma$ =0.2344, stepsize=100)
Loss	Cross Entropy Loss
dataset	cifar10
batch size	64
epoch	150

在による効果が区別できないので, 何本か適当な位置にショートカット設けたランダムアーキテクチャとの比較も行いたい.

## 4 今後の予定

- ランダムアーキテクチャとの比較
- DART の unrolling 実験

## 5 ソースコード

github の notebook リポジトリ参照.

## 1 今週やったこと

- 評価実験をたくさん

## 2 実験

10 回行った探索に対し各々評価を 1 回, VGG19 に対し異なるシード値で 10 回. それぞれ 10 回ずつ実験した.

表 1 に評価時の実験設定を示した. optuna によって得られた設定を利用した.

### 2.1 結果

評価時のグラフは./graph を参照. 表 2, 3 にはテスト精度の結果を示した.

## 3 考察

optuna で lr,  $\gamma$ , stepsize を最適化したが, 期待していた lr が 0.01, stepsize が 100 という値に近いパラメータで  $\gamma$  が得られた. train size を 20 分の 1 にしていても割とうまくいくのかもしれない.

google colab だと 3 時間かかっていたのが, usagi サーバーだと 1 時間早くなった. GPU の性能がクラウドの同期に時間がかかるのかは不明だが, google colab で開発して, サーバーで実験を回すのが捗るかもしれない.

ベースラインに対して有意な差があることが分かったが, DART による探索の効果とショートカットの存

表 2: 結果のテスト精度 (%)

	1	2	3	4	5	6	7	8	9	10
評価	93.95	94.09	93.32	93.58	93.66	93.65	93.66	93.52	93.80	93.76
ベースライン	92.97	92.95	93.25	92.90	93.06	92.93	93.07	93.07	93.03	93.06

表 3: 精度の比較

	test accracy mean $\pm$ std	delta
評価	93.6990 $\pm$ 0.2173	+ 0.6700
ベースライン	93.0290 $\pm$ 0.1002	