

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: LSTM の設定

component	LSTM + Linear
Optim	Adam(lr=0.01, $\beta=(0.5, 0.999)$)
Loss	MSE
hidden size	5 dim
data length	train=100, test=50
data size	train=10000, test=1000
epoch	250
batch size	100

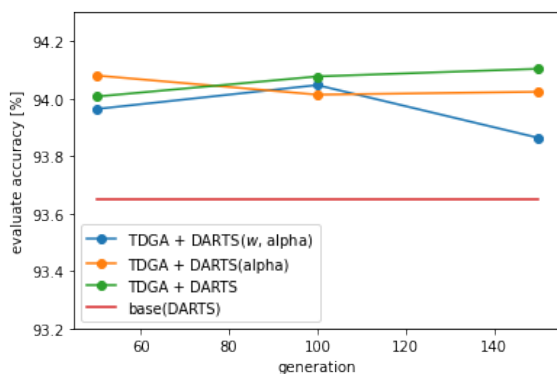


図 1: 3 回試行の結果

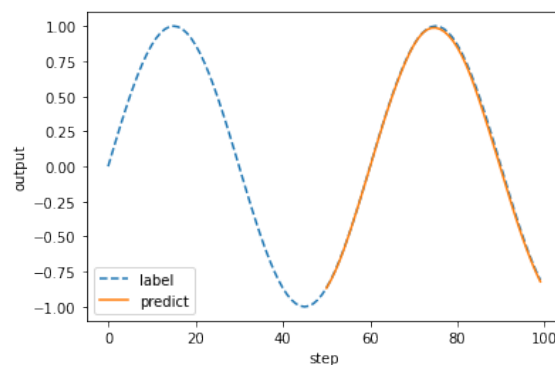


図 2: LSTM の動作確認 : sin 波

1 今週やったこと

- TDGA + DARTS の追加実験
- LSTM の動作確認

2 実験 1

TDGA + DARTS の試行回数が少なかったため, 追加実験を行った.

表 1 にモデルの設定を示す. 図 1 に合計 3 回試行した TDGA の実験結果を示した. 1 回試行のときと大きくは変わらず, TDGA のみによる探索が安定して, 精度の向上がみられる. しかし全体の改善は, 0.1 % 程度であり改善はなされていない.

3 実験 2

実験 1 のアーキテクチャ探索でネックとなる, アーキテクチャの学習や評価を RNN の利用によって, 処理を代替的に予測して, 探索時間を削減する実験の第一歩として, まず LSTM の知識や動作を確認した.

表 2 に LSTM の設定を示した. 時系列長 100 のデータから次の値を予測するように学習し, テストでは時系列長 50 のデータから再帰的に 50 ステップ先まで予測した.

図 2, 3 に sin 波などから作成したデータセットのテストの結果を示す. 値域が小さい周期関数は, ある程度高い精度で予測できることが分かった.

また Optimizer は SGD よりも Adam が格段に学習が早いので, Adam を採用する.

問題点

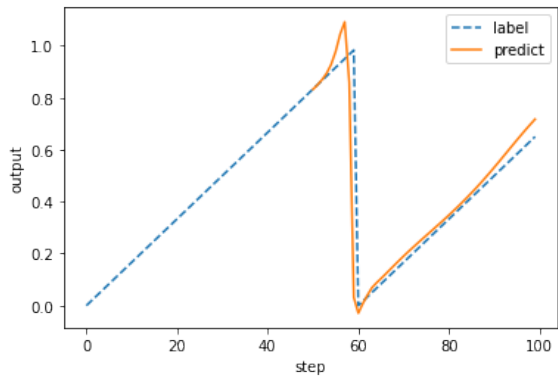


図 3: LSTM の動作確認：ノコギリ？

- $y = x$ など関数の値が大きすぎるとうまく学習できない (loss が不安定になるのが原因?)
- データ数が足りないので増やす

4 今後の予定

- データセットを作成する (データは現在増強中)
- LSTM で実験する