

進捗報告

1 今週やったこと

- GA の実験

2 GA の実験

ラベルなし画像 100 枚取り出し、それらに対するラベルを GA によって探索する。表 1, 2 に実験の設定を示す。遺伝子は 0 から 9 の整数値をとる整数値コーディングとした。

選択はサイズ 2 のトーナメント選択, 交叉には二点交叉, 突然変異は別の数値にランダムに移るように設定した。

また, 前回 train と eval の分け方を世代ごとに変えるようにした,

表 1: GA の設定

個体数	30
世代数	100
交叉率	1.0
突然変異率	0.02
全ラベル付き画像	250 枚
train:eval	100 枚:150 枚
search	100 枚

表 2: FixMatch の設定

model	WideResNet28-2	
data set	cifar10	
train data	labeled+serach	200
	unlabeled	49650
batch size	labeled+random	64
	unlabeled	64 * 7
val data	150	
num_iterations	5000	
optimizer	SGD(lr=0.1,momntum=0.9)	
loss	cross_entropy_loss	

2.1 結果

図 1 に示す。14 世代以降で fitness が落ちてしまっているのは train においてももとのラベル付きデータを入れ忘れてしまっていたからであった。ただし, それを抜きにしても精度の頭打ちが来ていることが分かる。

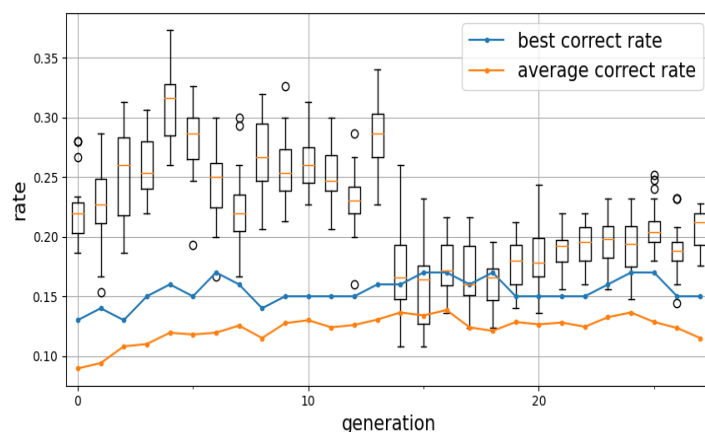


図 1: 結果

3 SSL の実験

あるラベルなしデータに各 0 から 9 までのラベルをつけて精度を求め, 最も高い精度のラベルをつけてラベル付きデータへと加えるような実験をした。表 3 に SVM の設定を示す。また, cifar10 使用し, ラ

表 3: SVM の設定

kernel	poly
C	3.654
γ	53.15
coef0	0.0
degree	3

ベル付きデータ 250 枚を train と val に 100 枚:150 枚で分けて探索する。

3.1 まとめ

時間がなく結果をまとめることができていないがポイントとしては以下のものが挙げられた。

- 大体 20%ほどの正答率であった。
- 正答ラベルの時だけがとびぬけて精度が良かったというわけではなかった。
- 正誤答関わらず初期の段階でつけられたラベルが同じものが続いたとき，そのあともそのラベルが多く出現する傾向にあった。

4 来週の課題

- 実験設定の改良