

## 進捗報告

### 1 今週やったこと

- GA の実験
- SSL の実験
- 論文の再調査

### 2 GA の実験

#### 2.1 実験 1

表 4, 5 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, 今回事前学習したモデルを初期モデルとして学習させた.

##### 2.1.1 結果

図 1, 2 に示す. 図 2 相関性はありそうではあるが, 学習が全くできていない. また 14 世代目が適応度

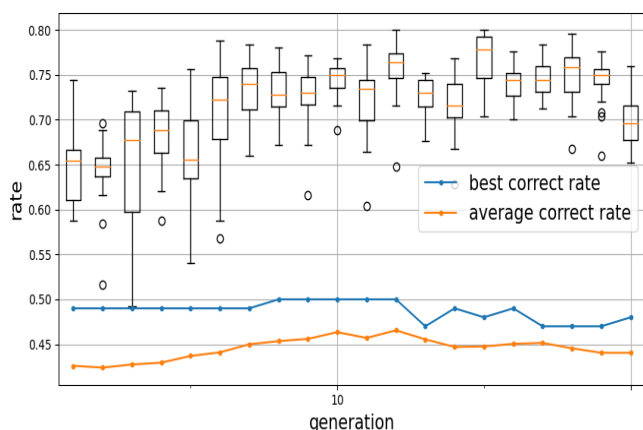


図 1: 実験 1 の結果

の平均が最高であり, その個体を使って再学習したものの精度を表 1 示す. 閾値はある遺伝子座の採択された遺伝子が全個体に対して占めていた割合に対するものである. 増やしたものがラベル付きデータ

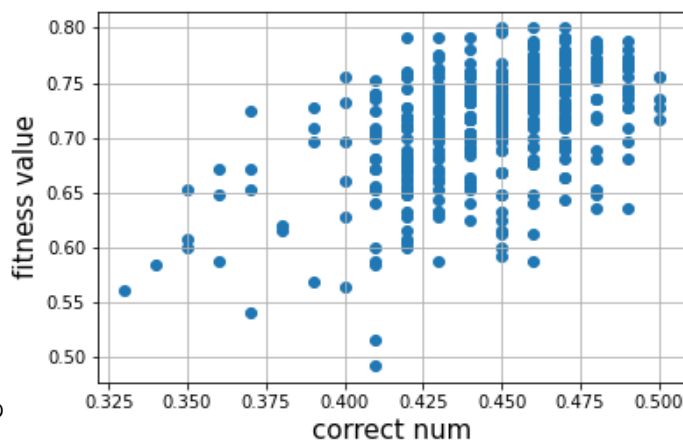


図 2: 実験 1 の相関図

表 1: GA の設定

採択数	正答数	閾値	精度
0			0.868
100	46	なし	0.836
39	22	0.19	0.862
19	14	0.2	0.825

のみを使った時よりも全て低い結果となった.

#### 2.2 実験 2

実験 1 を踏まえ, 個体評価において search 以外の unlabeled が邪魔しているのではと考え, 事前学習は FixMatch のまま個体評価に対して FixMatch ではなく cross entropy loss のみで学習した. 表 6, 7 にパラメータを示す.

##### 2.2.1 結果

図 3, 4 に示す. 図 4 から正答率が増えるほどばらつきがあり, これはデータ数が少ないことに起因しているのではないかと考えられる..

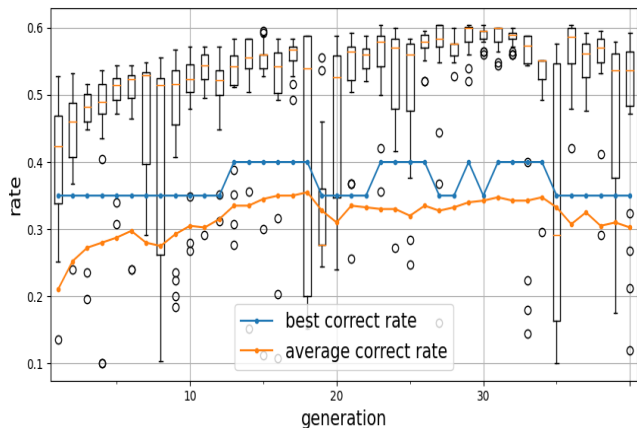


図 3: 実験 2 の結果

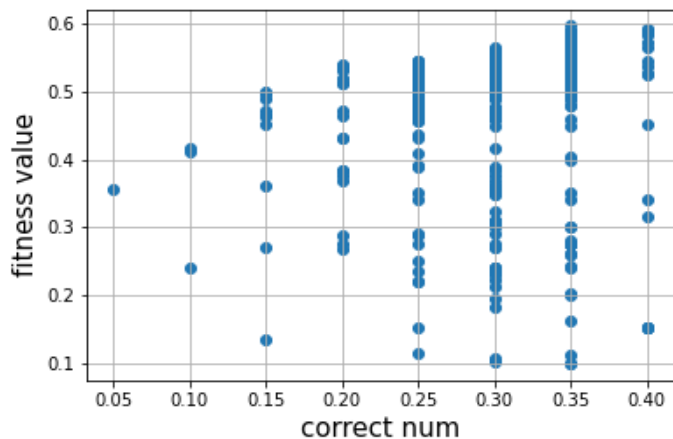


図 4: 実験 2 の相関図

### 3 SSL の実験

GA 以外を試す試みとして, FixMatch の loss に VAT の loss を入れて実験を行った. FixMatch の loss に対し 0.01 の重みを付加して VAT の loss を入れた.

表 3 に結果を示す. またこれは 2 回試行の平均である. 若干の改善が見れなくもないが検討の余地はあると思われる.

### 4 論文の再調査

GA を用いて半教師あり学習のラベル付きデータを増やす提案はされていたものの, 2 クラス分類問題に対するものがほとんどであり, cifar10 のような

表 2: 実験設定

model	WideResNet16-2	
data set	cifar10	
train data	labeled	100
	unlabeled	49900
batch size	labeled	32
	unlabeled	32 * 7
val data	10000	
num_iterations	2*16	
optimizer	SGD(lr=0.1,momntum=0.9)	

表 3: 結果

FixMatch	0.849
FixMatch+VAT	0.857

画像の分類問題の論文は調べた内にはなかった.

また, 最近の半教師あり学習は自己教師あり学習が組み込まれたものが精度を出している. 手法として以下のものがある.

- Ensemble AutoEncoder Transform(:EnAET)
- Semi-Supervised Self-Supervised Learning(:S4L)
- SimCLRv2
- CoMatch(MoCo)

また, vision transformer についても自己教師あり学習により精度の改善があったらしいのでそことも絡められる点でも GA よりも自己教師あり学習を絡めた研究をした方がよいかも知れないと思った.

### 5 来週の課題

- 実験設定の改良
- 自己教師あり学習の調査

表 4: GA の設定

個体数	20
世代数	20
交叉率	1.0
突然変異率	0.02
labeled	250 枚
search	100 枚

表 5: FixMatch の設定

model	WideResNet16-2	
data set	cifar10	
batch size	labeled	32
	unlabeled	32 * 7
optimizer	SGD(lr=0.1,momntum=0.9)	
loss	cross_entropy_loss	
事前学習		
train	labeled	100
	unlabeled	49650
val data	150	
num_iterations	2**15	
GA の評価		
train	search のみ	100
	unlabeled	49650
val data	250	
num_iterations	5000	
得られた個体の評価		
train	labeled+search	250+?
	unlabeled	49650
val data	10000	
num_iterations	2**16	

表 6: GA の設定

個体数	20
世代数	40
交叉率	1.0
突然変異率	0.02
labeled	100 枚
search	20 枚

表 7: FixMatch の設定

model	WideResNet16-2	
data set	cifar10	
batch size	labeled	32
	unlabeled	32 * 7
optimizer	SGD(lr=0.1,momntum=0.9)	
loss	cross_entropy_loss	
事前学習		
train	labeled	10
	unlabeled	49880
val data	90	
num_iterations	2**15	
GA の評価		
train	search のみ	20
	unlabeled	49880
val data	250	
num_iterations	5000	