

SIGNATE

【練習問題】 画像ラベリング (20種類)
結果報告

Team Green

Hidetake SUGO

m.Kiyo

Nishikawa

Ryo3 T(TeamShiny)

Rising Moon

0 : 本コンペのデータセットの事前調査

今回のコンペのデータセット $32 \times 32 \times 3 \text{ ch}$ のデータで他に使えるものはないか事前調査を行った。
同様のデータセットとしてCifar10及び/Cifar100があった



何故、外部のデータセットを調査するのか？

Cifar10等のデータセットはImagenetなどと同様にSOTA (state-of-the-art) なモデルの精度評価に使われている。
ここで高精度を出しているモデルを用いて、データセットを今回のデータセットに差し替えれば同様の精度が出せる事を期待した。

| データセット | クラス数 | 型式 | 訓練データ | テストデータ |
|----------|------|---|-------|--------|
| 本コンペ | 20 | $32 \times 32 \times 3 \text{ch}$ (カラー) | 5万枚 | 1万枚 |
| Cifar10 | 10 | $32 \times 32 \times 3 \text{ch}$ (カラー) | 5万枚 | 1万枚 |
| Cifar100 | 100 | $32 \times 32 \times 3 \text{ch}$ (カラー) | 5万枚 | 1万枚 |









クラス数以外はほぼ同じなのでそのまま使えそうな気がする。
懸念点としては、訓練データ数が同じなので、1クラス当たりのデータ量がCifar10の半分になるのでそれがどう精度に影響するか

1 : 最終的な解法以外に試して却下した解法とその理由

以下のサイトでCifar10で最高精度を出しているモデルが、EffNet-L2 (SAM)だった

<https://paperswithcode.com/sota/image-classification-on-cifar-10>

Cifar100の最高精度もEffNet-L2 (SAM)なので20クラスなら問題にならないと考えた。

| RANK | MODEL | PERCENTAGE CORRECT [↑] | PARAMS | EXTRA TRAINING DATA | PAPER | CODE | RESULT | YEAR |
|------|--------------------|------------------------------------|--------|---------------------------|--|---|---|------|
| 1 | EffNet-L2 (SAM) | 99.70 | | ✓ | Sharpness-Aware Minimization for Efficiently Improving Generalization |  |  | 2020 |
| 2 | ViT-H/14 | 99.5 | 632M | ✓ | An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale |  |  | 2020 |
| 3 | ViT-L/16 | 99.42 | 307M | ✓ | An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale |  |  | 2020 |
| 4 | BiT-L (ResNet) | 99.37 | | ✓ | Big Transfer (BiT): General Visual Representation Learning |  |  | 2019 |

SAMを用いて学習するも、うまく動かないor精度があまり出ない。

この時点での最高val_accは65%ぐらい。

引き続きEffNet-L2 (SAM)を使った手法をググって他のやり方を探す。

2 : 最終的に採用した解法とその理由 (その 1)

Effnet-L2(SAM)でググってたら、以下のページに出会う

「転移学習でCIFAR-10正解率99%を達成する方法」

<https://qiita.com/T-STAR/items/de933c67b52a47f9efef>

Colabリンク

<https://gist.github.com/T-STAR-LTD/65a4cfbb6d554e2c8309a0e>

本サイトを参考に以下の順序でEfficientNetB5で学習を行った

①対象クラス数に合わせたトップのみ学習(5 ep)



【以後、全層再学習】

②徐々に学習率を上昇(Warmup)させ再学習(10 ep)



③一定期間学習率をキープし再学習(10 ep)



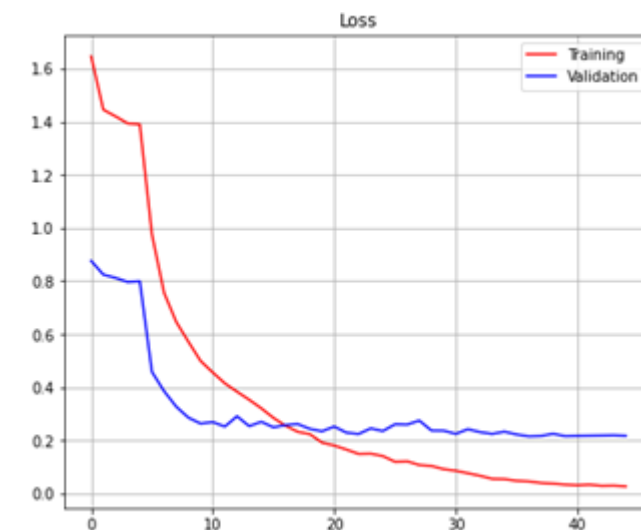
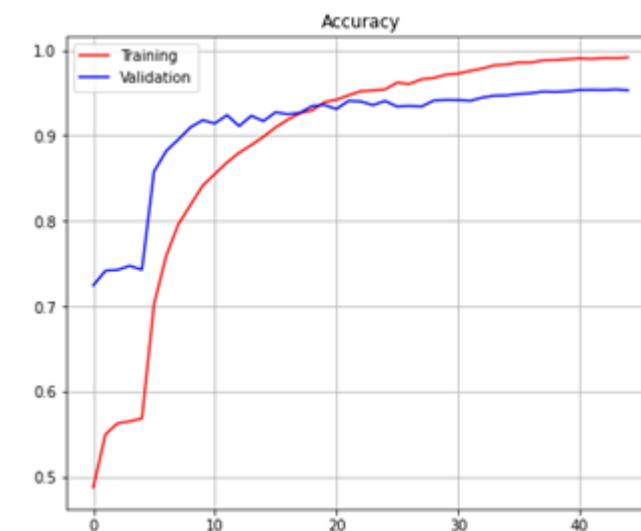
④徐々に学習率を減少(Cooldown)させて再学習(20 ep)



終了

上記の学習トータル45epで(Colab TPU環境で)
トップ1分×5ep=5分、残り3分×40ep=120分
で最終的なval_accは約95%に到達する。

EfficientNetB5での学習結果



2 : 最終的に採用した解法とその理由 (その 2)

前述のqiitaの作者がVision Transformerについても記事を掲載していたのでそれにもチャレンジしてみる

「Vision TransformerでCIFAR10正解率99%を達成する方法」

<https://qiita.com/T-STAR/items/728b5c279bcfb6a6381f>

Colabリンク

<https://gist.github.com/T-STAR-LTD/5c78f4a460541191a68708c29b>

Warmup : 4 ep

Cooldown : 16ep

上記の学習トータル20ep(Colab TPU環境で)

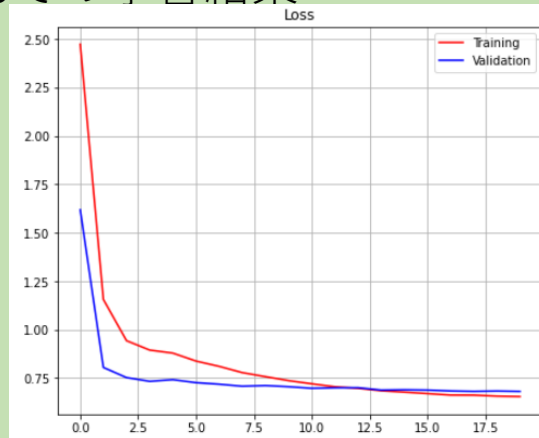
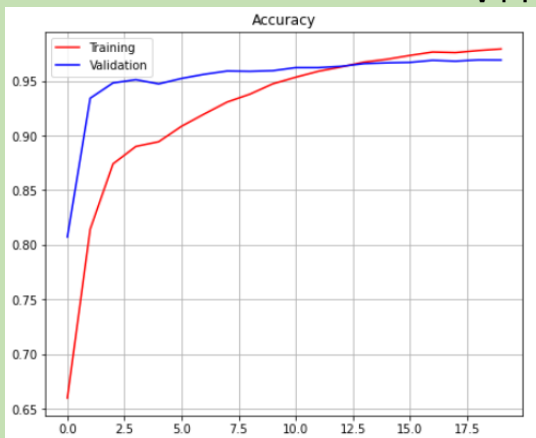
1分×20ep=20分 爆速で学習が進み

最終的なval_accは、

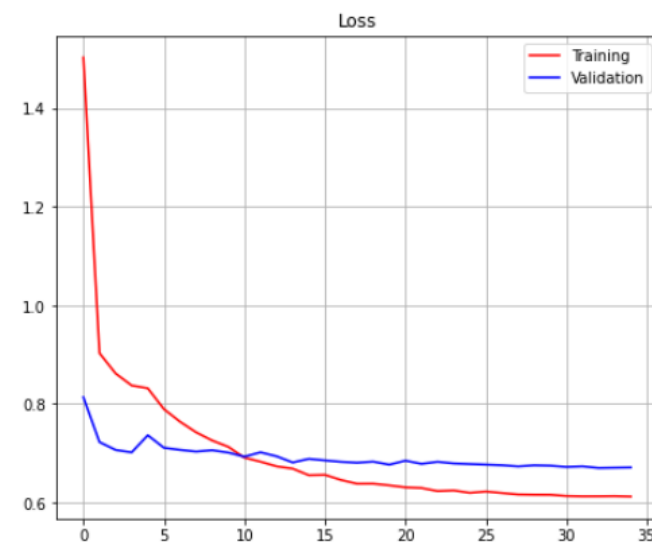
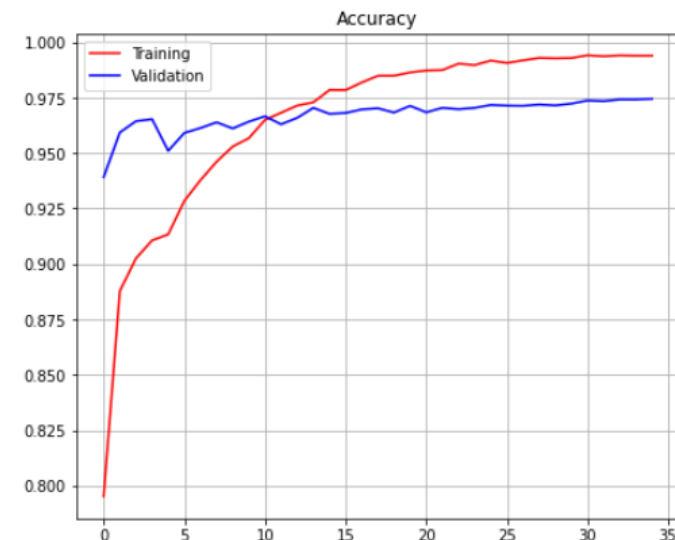
ViT-B_16では約96.9%に到達する。

ViT-L_16では35epで約97%に到達する

ViT-B_16での学習結果



ViT-L_16での学習結果



2：最終的に採用した解法とその理由（その3）

最終的には、上記その1 / その2の手法のアンサンブルを行い、最終的にチームとして1位を取得するに至った。

3：精度を上げるために工夫した事

○Cifer10などの既存のデータセットのSOTAをリサーチした。

○エポック数の調整

4：精度を上げる為に本当は試したかったが試せなかったこと

前述のようにSAMを使ったモデルがまともな精度が出せていないが、実質SOTAはSAMを使ったモデルであるので、SAMを使ってある程度高精度なモデルが作成出来たら、他のモデルとアンサンブルさせてどれだけの精度が出せるかを試してみたい。

Greenチームのソースコードは以下のGoogle Driveで公開予定です。

<https://drive.google.com/drive/folders/1J4f96r8jRQo6zmPkyqAjcsb4Ah4Z4cpo>

5：チームで効率よく作業するために工夫したこと

方針：既存/使い慣れたツールを使うことで、無理や無駄を極力なくす。



a. 連絡手段

Slackのグループチャットで完結

- ・任意の時間に確認可能、伝達漏れもなくす
- ・特に集合等はせずに、各自の本業の状況に合わせることが可能



b. コード/データ共有方法

google driveの共有機能を使用(colabと好相性)

pickleで前処理をしたデータをチーム内で共有

※Githubも検討したが、今回はver管理までは不要と判断。



c. 実施環境

Google Colab

- ・メンバのローカル環境に依存せず、コードを回せる
(ローカルを挑戦したが、ver違い等により煩雑になってしまう)
 - ・物体検出だと、GPU/メモリが必須。無償でそれなりのリソースを確保できる。
- ※前処理等、部分的にはローカル環境との使い分けも有り。

6：チームで効率よく作業をするために反省していること

スコアは1位であったが、協業(ギグワーククエスト)という点では課題点が明らかになった。

a. マネジメント面(プロマネの不在)

- ・方針立案…スコア上昇への戦略の妥当性/網羅性
- ・チーム編成/役割分担…誰が何をするのか、各自の強みを活かすには？
- ・タイムマネジメント…いつまでに何をどこまでやるのか。
- ・連絡伝達手段…Slackのみが最適かどうかはわからない。

※ギグワークとなると各メンバーが対等の関係となるため、(一般的な)管理職的立ち位置が自然発生しづらい。
だからこそ、事前にリーダーの位置づけになる人を決定しておく必要がある。

b. スキル面

- ・レベルの全体的底上げ…エース1人に依存する協業は安定した成果につながりづらい
理想は誰が抜けても、チームで穴を埋めれるくらいの技術力/連携が必要か。

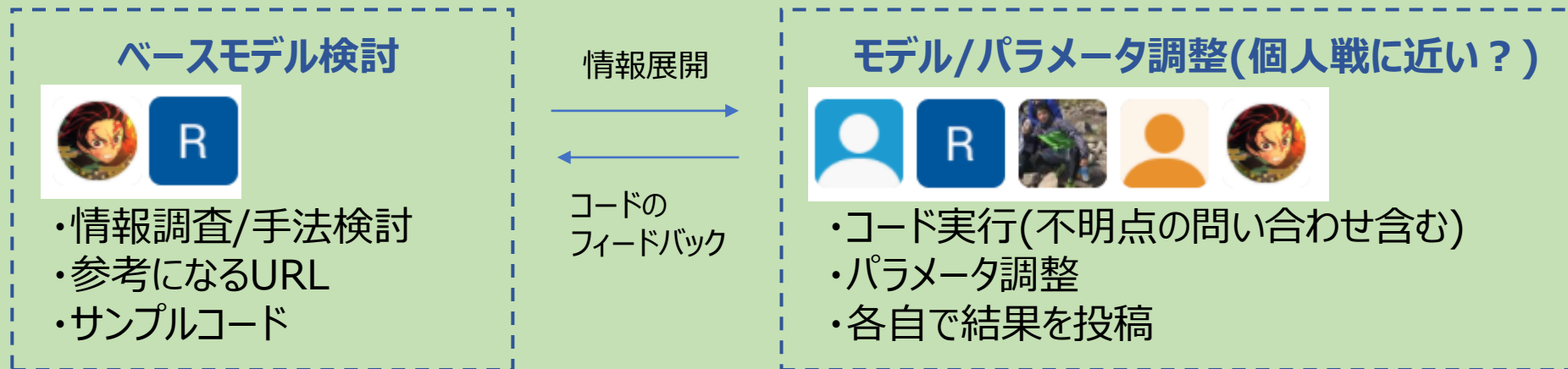
※勉強会や情報共有会などの継続的实施で日々能力を研鑽することが重要

7 : お題発表までの準備期間中に実施したこと（上記回答と重複してよい）

- ・連絡方法/コード共有/開発環境の統一
- ・事前の情報収集(各自)
- ・各PBLの上位ランキング者のコード確認(各自) - 構造化データ、物体検出 etc.

8 : 各役割の大雑把な役割分担

特に細かく決定してなかったが、進めていくうえで以下のような役割に落ち着いた。



9 : 事前に立てていた戦略と、実際とのギャップ/気づき

役割は最初から割り当てられていたが、お題を見てから細かい役割を決めようとしていた。

→ **お題の特性※上、役割を明確化する必要性が乏しく突っ走ることになってしまった。**

(基本的なことかもしれないが)異種のモデルを組み合わせると、例えば個々のスコアが良くなくとも、アンサンブルした際のスコア改善の効果が大きい。

→ **いろいろなモデルを試すことは決して無駄にならない**

何よりも最初の情報収集(Qiita記事、英語記事、中国語も?)がすごい大事

→ **常日頃 + お題が決まってから の情報収集力**

提案：次回以降、チームとしてスコア向上に取り組むのであれば、

- ① 目ぼしい手法を短時間で調べていったんすべて列挙する。
- ② その列挙した手法ごとに、チーム内で担当を決める。
- ③ スコアが上がるように調整した結果を持ち寄ってアンサンブル実施
- ④ ②～③を短いサイクルでコンペ期間中に何回か回す

特に、今回は**解放立案士**と**品質保証士 (プロマネ)**の役割が重要



解放立案士

プロジェクト初期に
手法の網羅的列挙



品質保証士(プロマネ)

進捗管理、
チームの役割調整



その他、職種

各自の基本スキルを向上させ、
チームの誰でも対応可能とする