

進捗報告

1 今週やったこと

- レポ移行, cifer-10 の論文ベースラインを確認
- TDGA AutoAugment 実験

2 レポ移行

FastAutoAugment のレポから移行完了。また、このタイミングで用いているモデル及び最適化スケジューラ等を一通り見直し、RandAugment で挙げられていた cifer-10 の自動拡張なしベースラインを達成できることを確認した。CNN 単体の学習は Wide-ResNet-28-2, 200 エポックで 50 分 (accuracy:95.1%), Wide-ResNet-28-10 で 6 時間 (accuracy:96.2%) ほど。また、cudnn.benchmark を設定することで元のレポの 3 時間ちょっとかかっていたものが 2 時間程度で完了するようになった。PyTorch を使って入力画像サイズとネットワークが変化しないようなモデルを使っている人は cudnn.benchmark=True 設定を使ってみると良いかも。

3 TDGA AutoAugment 実験

新しいモデルを用いて温度パラメータ、強度を変化させつつ実験した。温度パラメータは [0.01, 0.08] を 0.01 刻み、強度は [0, 30] を 5 刻みで回した。また、温度パラメータは探索中固定である。

cifer-10 のパラメータは 200 エポック、optimizer は SGD、スケジューラはコサインアニーリング (PyTorch の torch.optim.lr_scheduler.CosineAnnealingLR を参照) を用いている。これらの値は RandAugment の論文で採用されているものである。またデフォルト拡張としては pad-and-crop、ランダムフリップ、および画素値の正規化を用いている。

図 1 に強度を 15 に固定して温度を変化させた Test accuracy を示す。

図 2 に温度を 0.05 に固定して強度を変化させた Test accuracy を示す。

どちらも 1σ の範囲内という感じでそこまでの精度向上は見られない。選ばれている拡張も先週の報告にあるような画像に変形を加えないようなものが主に選ばれていた。

ここで、図 3 に RandAugment で報告されていた、各操作を採用することによる精度向上の表を示す (正確には表ですがスクショしたので図になっています)。△はその操作を選択することで最終精度が向上する期待値を表している。

これを見るとやはり変形を加えるようなものの方が精度への寄与は大きいことがわかる。このあたりの多様度をどのように維持するのかということが今後の主な課題になりそう。

表 1 に他手法との比較を示す。比較に用いた手法は Population Based Augmentation (PBA) [1], Fast AutoAugment (Fast AA) [2], AutoAugment (AA) [3] および RandAugment (RA) [4] である。TDGA Augment 以外の値は RandAugment で掲載されているものである。また、baseline の値は手元で確認した 10 回試行の平均であり、() 内の値は論文で挙げられていたものである。また、Wide-ResNet-28-10 を用いた実験は時間がかかるためまずは Wide-ResNet-28-2 でチューニングする。

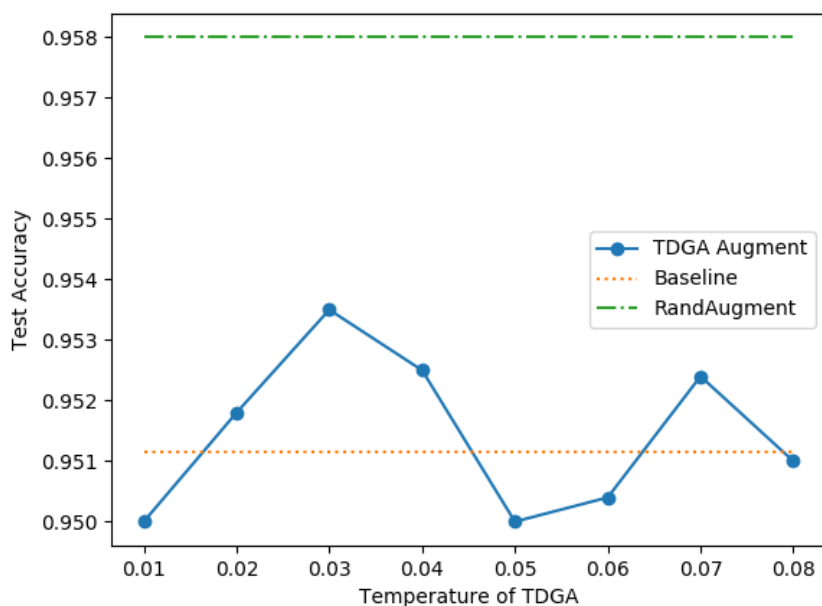


図 1: 温度を変化させた Test accuracy(強度固定)

表 1: Test accuracy(%) on CIFER-10

	baseline	PBA	Fast AA	AA	RA	TDGA Augment
Wide-ResNet-28-2	0.951(94.9)	-	-	95.9	95.8	0.9535
Wide-ResNet-28-10	-(96.1)	97.4	97.3	97.4	97.3	-

4 強度をランダムにした RandAugment について

先週操作の強度を一定にしているのかという議論があったが, RandAugment の付録 A.1.1, A.1.2 に書いてあった. A.1.1 によると, 強度 M をランダムに選ぶ手法, 一定の強度で統一する手法 (論文内容), 線形に増加させる手法, 上限を徐々に増加させるランダム選択などを試した結果そこまで精度に変わりがなく, パラメータとして単純な一定の強度を採用したという話.

また, A.1.2 によると, 一定の値から 1 つの操作の強度だけを変化させることによる影響が小さいという結果から, 探索空間の削減を考えて一定の強度を採用するに至ったらしい.

5 来週のタスク

操作の種類毎に多様性をもたせるように拡張したい.

6 TODO タスク

- 他のモデルを用いた実験
- CIFER-100, SVHN など他ベンチマーク問題に対する精度比較
- 漫画画像への適用

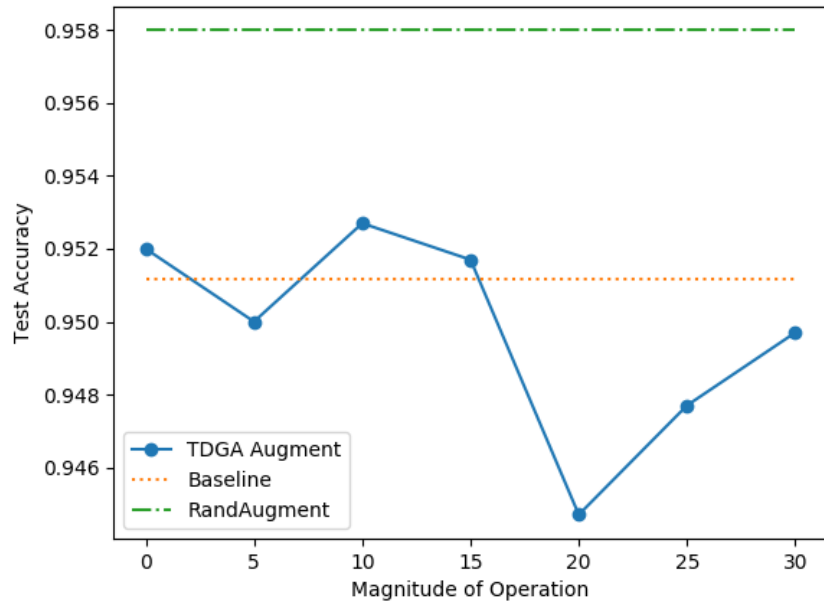


図 2: 強度を変化させた Test accuracy(温度固定)

参考文献

- [1] Daniel Ho, Eric Liang, Ion Stoica, Pieter Abbeel, and Xi Chen. Population based augmentation: Efficient learning of augmentation policy schedules. *CoRR*, abs/1905.05393, 2019.
- [2] Sungbin Lim, Ildoo Kim, Taesup Kim, Chiheon Kim, and Sungwoong Kim. Fast autoaugment. *CoRR*, abs/1905.00397, 2019.
- [3] Hiroshi Inoue. Data augmentation by pairing samples for images classification. *CoRR*, abs/1801.02929, 2018.
- [4] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V. Le. Randaugment: Practical automated data augmentation with a reduced search space. *arXiv: Computer Vision and Pattern Recognition*, 2019.

transformation	Δ (%)	transformation	Δ (%)
rotate	1.3	shear-x	0.9
shear-y	0.9	translate-y	0.4
translate-x	0.4	autoContrast	0.1
sharpness	0.1	identity	0.1
contrast	0.0	color	0.0
brightness	0.0	equalize	-0.0
solarize	-0.1	posterize	-0.3

図 3: 各操作を選択することによる精度への影響