

Deep Learning 向け Data Augmentation の評価手法の提案 -Fréchet Inception Distance に基づく方法-

小林 賢一

辻 順平

能登 正人

神奈川大学大学院工学研究科電気電子情報工学専攻

1 はじめに

Deep Learning には大量の学習データが必要であり、データの収集が困難な場合においては、data augmentation (DA) という手法が提案されている。DA には、データセットごとに適したパラメータや手法の組み合わせ法があることが知られているが、それらを効果的に見つけるための手法は知られていない。

そこで本研究では、DA により拡張した学習データの良し悪しを評価する指標に Fréchet Inception Distance (FID) を用いることを提案する。実験結果から、FID を用いた学習データの分布と検証データの分布の間の距離の大小と正答率の間に相関があることを明らかにした。したがって、この指標を用いることで、データセットごとに適したパラメータや手法の組み合わせ法を効果的に見つけることが可能であることが示唆された。

2 FID に基づく DA の評価

機械学習の学習データと検証データの一般論によると学習において良い性能を示すためには、学習データと検証データの差を小さくする必要があると言われていた。この知見を DA に適用すると DA によって学習データの分布を検証データの分布に近づけることができれば正答率が向上すると期待できる。このことから、DA を適用した学習データの分布と検証データの分布の間の距離を測り、測定した距離の度合いに応じて正答率が向上または低下するのかを事前に把握することが可能であると期待できる。

このような背景から我々は、学習データの分布と検証データの分布の間の距離を測る手法として FID を用いる [1]。理由は、FID は入力する画像の CNN での分類のしやすさに応じて値を示す指標となっている。し

たがって FID により測定する分布間の距離は、正答率と相関があることが想定されるためである。

そこで本研究では、FID により導き出す分布間の距離は DA 利用における困難性を解消することが可能であるか検討を行う (本研究の概要図 1)。このためには、まず FID と正答率相関関係を明らかにする必要があるため、FID と正答率の相関係数を求める実験を行う。

学習データの分布と検証データの分布の間の距離を FID を算出する方法を示す。画像の集合を A_i 、その要素を $a \in A_i$ とする。ここで Inception-v3 モデルに対し a を入力して得られる 2048 次元のベクトル $h(a)$ を求める。 $h(a)$ の分布が多変量正規分布に従うと仮定し、2つの分布を求めると、その分布間の Fréchet Distance を計算ができる。 A_i に対応する平均ベクトル μ_i と共分散行列 Σ_i は、次式で計算できる。

$$\mu_i = \frac{1}{|A_i|} \sum_{a \in A_i} h(a)$$

$$\Sigma_i = \frac{1}{|A_i| - 1} \sum_{a \in A_i} (h(a) - \mu_i)(h(a) - \mu_i)^T$$

これらを用いて 2つの画像集合 A_1 と A_2 において FID を求める。

$$\text{FID}(A_1, A_2) = \|\mu_1 - \mu_2\|^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1 \times \Sigma_2))^{\frac{1}{2}}$$

3 FID と正答率の相関関係の分析

学習データと検証データの集合間の FID と正答率の相関係数を求め、2つの相関関係を明らかにする実験について述べる。

CNN の研究で用られる MNIST (M) 及び CIFAR-10 (C10), CIFAR-100 (C100) のデータセットを用い実験を行う。また、C10 および M のデータセットは、DA を用いた場合の精度の変化が小さいため、データの元の量を半分に減少させたデータセット (C10 $\frac{1}{2}$, M $\frac{1}{2}$) も

Evaluation of Data Augmentation for Deep Learning -Evaluation Method Based on Fréchet Inception Distance-

Kenichi Kobayashi, Junpei Tsuji and Masato Noto
Graduate School of Electrical, Electronics and Information Engineering, Kanagawa University

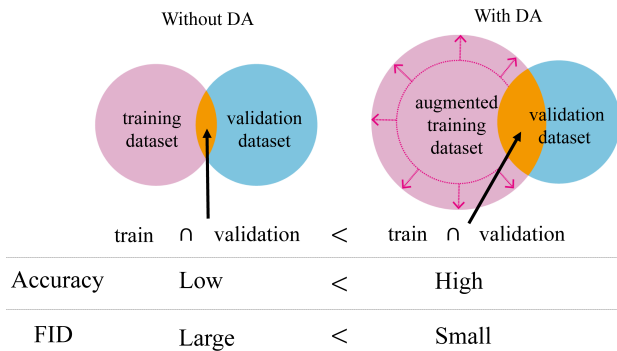


図 1: 本研究の概要図

本研究では用いる。したがって、実験のデータセットの種類は5種類となる。各データセットは Holdout 交差検証を行い、ランダムな学習データと検証データの組み合わせを3種類作成する。使用する CNN モデルは、全て Deep Learning 用ライブラリ Keras 上の提供モデルを参考に構築している¹。

DA 手法は、rotation (ro), zoom (zo), crop (cr), shear (sh), cutout (cu), mirror (mi) の6種類を用いる。DA の適用は、機械学習用 DA ライブラリである Augmentor を用いる²。各 DA 手法のパラメータの設定値は、各 DA 手法ごとに Augmentor 上のパラメータの設定可能な範囲を5等分し、それぞれにおいて DA 適用を行う。

本実験では、DA 手法を単体で適用した場合および複数の DA 手法を組み合わせで適用した場合の2種類について実験を行う。前者では、mi 以外の5つの DA 手法をそれぞれ適用した場合の実験を行う。後者では、ro および cr および mi を組み合わせた3つの適用方法においての実験を行う。

実験は、次の流れで行う。まず、各データセットにおいて各実験条件に従い、DA により学習データの拡張を行う。次に、拡張された学習データと検証データを用いて正答率および FID の算出を行う。最後に各実験条件ごとの正答率と FID を用い相関係数を求める。

4 実験結果

5 種類の DA 手法をそれぞれ単体で適用した場合および複数の DA 手法を組み合わせで適用した場合における正答率と FID の相関係数の表 1, 2 に示す。相関係数の値は、最大で C10 $\frac{1}{2}$ の zoom の場合に -0.96 、最小で M $\frac{1}{2}$ の zoom の場合に -0.20 という結果であった。

¹Keras, <https://github.com/keras-team/keras-docs-ja>

²Augmentor, <http://augmentor.readthedocs.io/>

また、M は他のデータセットよりも相関係数が低い傾向があることがわかる。全ての条件で、FID と正答率には負の相関があることが確認された。

表 1: DA 手法を単体で適用した場合の相関係数

	M	C10	C100	M $\frac{1}{2}$	C10 $\frac{1}{2}$
cu	-0.42	-0.56	-0.78	-0.82	-0.93
ro	-0.53	-0.77	-0.43	-0.50	-0.34
sh	-0.59	-0.85	-0.61	-0.92	-0.34
zo	-0.38	-0.85	-0.86	-0.20	-0.96
cr	-0.56	-0.83	-0.80	-0.82	-0.93

表 2: 2 種の DA を組み合わせた場合の相関係数

	M	C10	C100
ro & mi	-0.69	-0.75	-0.79
ro & cr	-0.89	-0.79	-0.86
cr & mi	-0.37	-0.86	-0.85

5 おわりに

本研究では、DA により拡張した学習データの良し悪しを評価するために、FID に基づく学習データの分布と検証データの分布の間の距離の大きさによる評価指標を提案した。

FID と正答率の相関分析の結果、学習データの集合と検証データの集合の間の FID を測定し、FID の大小と正答率の間に負の相関があり、FID の距離が小さくなるほど正答率は向上することが明らかとなった。この評価指標を用いることによって、FID の値に基づきパラメータや手法の組み合わせを効果的に見つけることが可能であると期待できる。

参考文献

- [1] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B. and Hochreiter, S.: GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium, *Proceedings of the Advances in Neural Information Processing Systems 30*, pp. 6629–6640 (2017).