

CNN による画像認識精度向上のための付加ノイズの検討

An examination of additional noise for improving image recognition using CNN

野田 遼太郎[†]
Ryotaro Noda

今井 信太郎[†]
Shintaro Imai

武田 敦志[‡]
Atsushi Takeda

1 はじめに

近年、ニューラルネットワークを用いた画像認識に注目が集まっており、研究が盛んに行われている。中でも畳み込みニューラルネットワーク (Convolutional Neural Network, 以下 CNN) は現在も精度の向上が続いており、今後のニューラルネットワークを用いた画像認識で中心となるであろう手法である。CNN を用いた画像認識では、後述する Dropout (2.1 節) や Shake (2.2 節) などといった、学習途中のニューロンに影響を与える方法が精度向上の要因になっていることがわかっている。そこで、我々の研究グループでは、学習時に複数回ノイズを付加する手法を提案している。しかし、どのような、そしてどのようにノイズを付加すると効果的であるかは明らかでない。よって本研究では、より画像認識の精度向上に効果的な付加ノイズを検討する。

2 関連研究

2.1 Dropout

汎用的なニューラルネットワークモデルを作る方法として、Dropout が知られている [1]。Dropout は学習ごとにニューロンをランダムに消去する手法で、テスト時には消去した割合を乗算して出力する。Dropout を適用したニューラルネットワークはアンサンブル学習と同様の効果により過学習が抑えられ、学習の停滞が起りづらくなる。

2.2 Shake-Shake Regularization

2017 年 2 月時点で CIFAR10 の SOTA(State of the art) となった Shake-Shake regularization[2] では、Forward と Backward のパスをランダムに変える手法を使っている。特徴として、他 CNN と比べエラー率の停滞が少なく下がり続ける傾向にあることが挙げられる。

3 NoiseNet

本研究では、CNN の学習時にランダムにノイズを付加することにより学習の停滞を抑え、テスト時の認識精度向上を目指す。

提案手法のネットワーク構成を図 1 および図 2 に示す。本提案手法は、現在 CNN で主流となっている ResNet[3] を元にした構成にしており、ResNet における残差ブロックにあたる block の最後にノイズを付加する層の noise を配置している。図 1 は、ブロック数を仮に 3 とした場合の構成である。block 内の処理は図 2 に示すように、

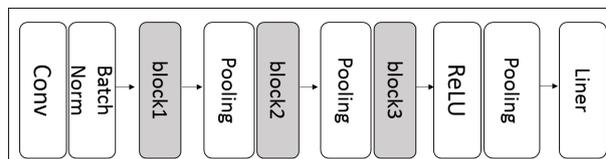


図 1: NoiseNet の構成

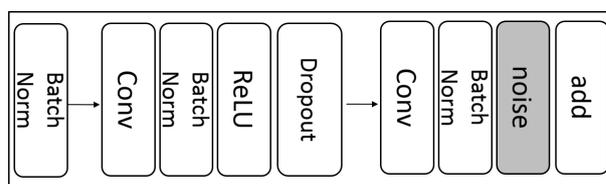


図 2: block の構成

学習速度を早めるための Batch Normalization や、特徴を抽出する畳み込み層、勾配消失を抑える Dropout などの層が配置されており、最後にノイズを付加する層が入っている。今回検討するものは、図 2 中の noise で乗算するノイズである。検討する項目については、付加するノイズの種類、ノイズを付加する回数、ノイズのパラメータとする。

4 実験

4.1 実験環境

本実験では、画像認識分野において有名なデータセットの一つである CIFAR-100 を用いる。実験環境を以下に示す。

OS: ArchLinux

GPU: GeForce GTX 970

メモリ: 16GB

言語: Python3

ライブラリ: chainer

4.2 付加するノイズの種類

本手法に使用するノイズの種類を検討するため、ガウス分布と一様分布の 2 種類の方式に従い生成した数を用いたノイズを比較する。ガウス分布は平均 1、標準偏差 0.2 で生成し、一様分布は平均 1、最小値 0.8、最大値 1.2 で生成した数を使用した。実験結果を図 3 に示す。一様分布を使用した場合は 300epoch 時点で loss が約 1.321、ガウス分布を使用した場合は 300epoch 時点で loss が約 1.154 となり、ガウス分布を用いたネットワークの方が精度が高いという結果になった。

4.3 block 数

図 1, 2 で示したノイズ付加を含む block 数を 2, 3, 4 に変更したものを用意した。ノイズには 4.2 節の結果

[†] 岩手県立大学, Iwate Prefectural University

[‡] 東北学院大学, Tohoku Gakuin University

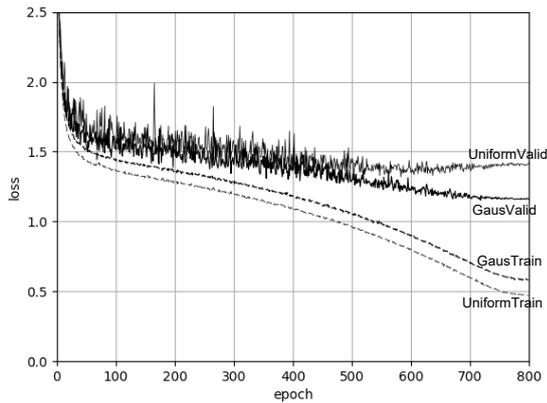


図3: ガウス分布と一様分布の比較

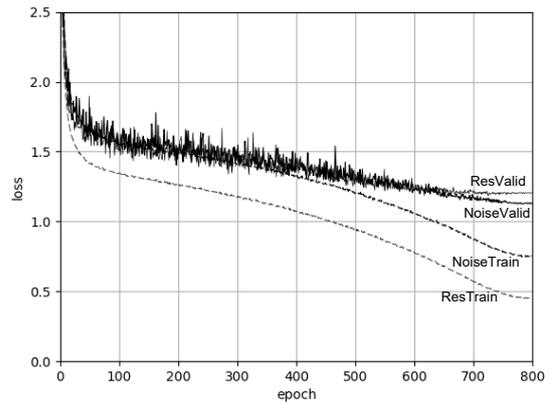


図5: 提案手法と ResNet の比較

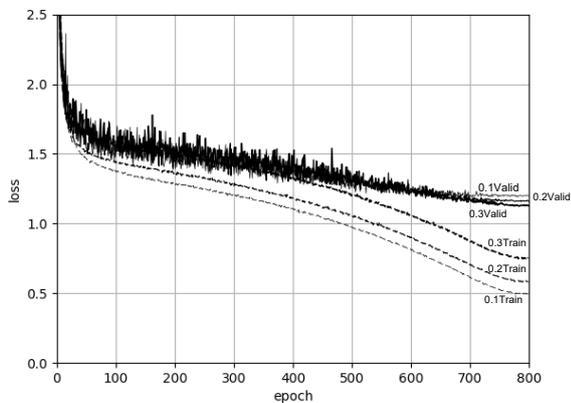


図4: ノイズのパラメータの比較

からガウス分布を使用し、同様のパラメータを使用する。block 数 2 と 3 を比較すると、最終的な精度はほぼ同じとなったが、block 数 2 の場合は 250epoch までの収束が若干遅く、block 数 3 と 4 を比較すると、テスト時の loss の差がそれぞれ 0.476, 1.224 となり、block 数 4 のものは過学習ぎみになってしまった。以上より、block 数は 3 が最適と言える。

4.4 ノイズのパラメータ

ノイズ生成に使用するガウス分布の標準偏差の値を変更し実験を行った。実験結果を図 4 に示す。標準偏差 0.1, 0.2, 0.3 の結果はそれぞれ loss 1.197, 1.160, 1.127 であり、0.3 の認識精度が最も高かった。また、0.5 では学習時とテスト時の認識率の逆転がみられ、0.7 では loss が発散してしまった。なお、モデルの性質上、大きな値や負の値が与えられると発散する可能性が大きく、効果が薄いことが推測されるため、平均値に関しては変更しないものとする。

4.5 評価実験

4.2, 4.3, 4.4 節の結果から、もっとも良い結果となったネットワーク構成である block 数 3 で平均 1, 標準偏差 0.3 のガウス分布を用いたノイズを加えた提案手法とノイズを加えていない ResNet を比較した。実験結果

を図 5 に示す。図 5 から、提案手法による学習が通常の ResNet に比べ loss が低いことがわかる。よってノイズを乗算する本手法は、有効であると言える。

5 おわりに

本稿では、学習時に複数回ノイズを付加して画像認識精度を向上させる手法において CNN に付加するノイズを検討し、付加ノイズには平均 1 標準偏差 0.3 のガウス分布を使用した block 数 3 のノイズを付加することで精度の高いネットワークになることがわかった。また、CNN の精度向上によりノイズ付加の有効性を示した。

今後の課題として、最適なネットワーク構成の検討が挙げられる。本実験では、ResNet をベースとした NN に対しノイズを付加することで精度向上を目指したが、ネットワークの構成に関してはほとんど触れていない。より精度向上に効果的なネットワーク構成およびノイズ付加の位置を検討する必要がある。また、ResNet 以外の NN に対しても同様にノイズ付加が効果的かどうかの調査も今後の課題とする。

参考文献

- [1] 斎藤康毅, “ゼロから作る Deep Learning,” オライリー・ジャパン (2016).
- [2] Xavier Gastaldi, “Shake-Shake regularization,” arXiv preprint arXiv:1705.07485 (2017).
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, “Deep Residual Learning for Image Recognition,” arXiv preprint arXiv:1512.03385 (2015).