

IoT を想定とする AutoEncoder と画像認識を 組み合わせた深層学習モデルの設計と検証

1240156 明神 拓歩 (Soft Intelligent System on a Chip 研究室)
(指導教員 星野 孝総 准教授)

1. はじめに

第3次 AI ブームでは、機械学習が急速に進化した。自己学習可能な AI が台頭し、特に画像認識分野では ILSVRC 大会がきっかけとなり、2012 年の AlexNet から始まる深層学習モデルが登場した。顔認識や不良品検品などに広く活用され、同時に AI ブーム以前の大量の行列計算が必要なニューラルネットワークも進化し、AutoEncoder などが生まれた。これらのモデルには CNN と Pooling が主に使われ、畳み込みとプーリングを組み合わせて特徴を抽出する基本構造である CNN が一般的になった。また、AI と IoT の連携が進み、IoT デバイスが生成するデータを AI が活用し、革新的な機能を提供した。これにより、現代の課題に対応し、セキュリティや効率的なソリューションが可能になった。

本研究では、AutoEncoder と画像認識を組み合わせた深層学習モデルの性能評価実験を行った。

2. 実験手法と内容

実験では、cifar-10 データセットを使って、画像を 3 クラスにまで減らし、 32×32 から 256×256 に広げたデータを使用した。使用言語は Python である。

実験では、下の図 1 のように AutoEncoder を使って画像を異なるサイズまで圧縮し、再構築する構成にした。畳み込みには 7×7 のカーネルを使い、各サイズに異なる層数を試した。また、最大値プーリングを使用して比較した。最適化には Adam を使い、画像認識と画像復元の実験も行った。実験から得られた結果では、AutoEncoder で学習時と評価時の損失を抑えることが重要で、画像復元と画像認識の両方で性能を確認した。

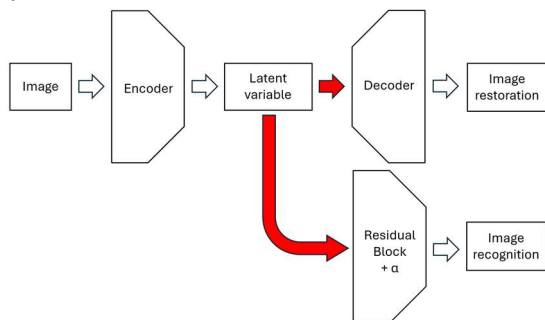


図 1: AutoEncoder と画像認識を組み合わせた
深層学習モデルの概要

3. 実験結果

本実験では、最大値プーリングにおける AutoEncoder と画像認識の結果について述べる。

AutoEncoder の実験結果を図 2 と表 1 に示す。図 2 のような 256×256 サイズから 208×208 まで圧縮する AutoEncoder が人の目でも画像認識できる基準で画像復元できた。表 1 に、実際に復元したすべての画像における各チャンネルの類似度を示す。ほとんどのモデルにおいて現画像と復元画像の類似度が 0.97 以上であった。図 2 は、実際に画像復元を行った図である。一番左の画像が入力画像であり、中央の画像が Encoder の出力画像を同じサイズに拡大したもの、一番右側の画像が復元された画像である。大まかな特徴を維持したまま学習できていることが確認できた。

画像認識の結果は図 3 に示す。図 3 では、過学習にならず

に Train Accuracy が 90%以上、Test Accuracy が 85%以上となった。潜在変数の画像がセグメンテーション画像を入力として、この精度が算出されたことは、適切にモデルが学習できたと考える。他の違うサイズの潜在変数における精度も同様に過学習にはならなかったが画像を圧縮するごとに精度が落ちていたことが確認できた。

表 1. AutoEncoder における画像の類似度

Correlation / image size	196×196	202×202	208×208	214×214	220×220
Red	0.95	0.97	0.97	0.97	0.97
Green	0.96	0.97	0.97	0.97	0.97
Blue	0.01	0.96	0.96	0.97	0.97
Average	0.64	0.97	0.97	0.97	0.97

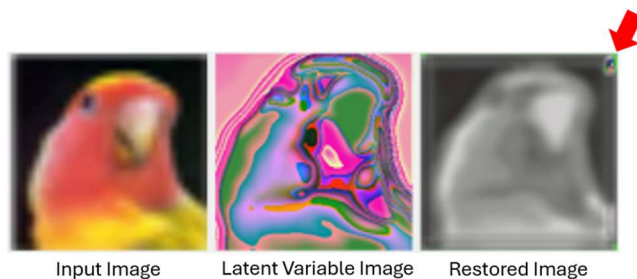


図 2. 潜在変数の出力サイズが 208 における
AutoEncoder の復元過程

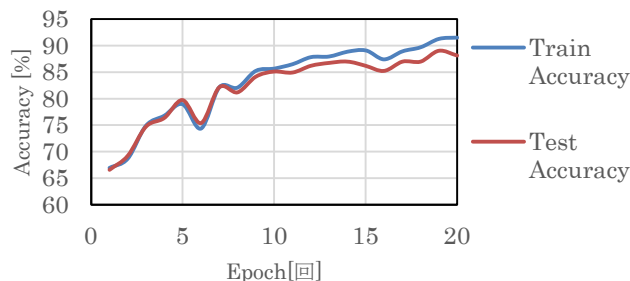


図 3. 潜在変数の出力サイズが 208 における画像認識の
1 学習あたりの Accuracy の変化

4. おわりに

まず、AutoEncoder では復元画像が全てのモデルで白黒画像になっていた。そのため、カラー画像になるように正規化の範囲を見直す必要がある。また、復元画像にノイズが発生したため、その解決にも取り組む必要がある。

画像認識では、入力画像の結果に近づけるように工夫する必要がある。さらに、画像認識結果のみでは AI の説明可能性が全くないため、その結果と理由について確認する必要がある。

参考文献

- [1] Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *science* 313.5786 (2006): 504-507.
- [2] LeCun, Yann, et al. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998): 2278-2324.