

畳み込みオートエンコーダによる次元削減を用いた画像クラスタリング

Image Clustering using Dimensionality Reduction by Convolutional Autoencoder

長谷川 海太† 黒木 啓之†

Kanata HASEGAWA† Takashi KUROKI†

† 東京都立産業技術高等専門学校

†Tokyo Metropolitan College of Industrial Technology

1 目的

機械学習において、音声や画像のような高次元のデータが扱われることは非常に多い。このようなデータに対し、データの次元数が大きいとそれに伴い計算量が増加してしまう「次元の呪い」と呼ばれる現象が発生することがある。この現象を回避するために、特徴を保持したままより少ない次元数で表現する手法として次元削減がある。

次元削減を行う手法の一つにオートエンコーダがある。これを用いる手法として、SDA(Stacked Denoising Autoencoder)で次元削減を行い、クラスタリングを行う手法が提案されている [1]。

本研究では、畳み込みオートエンコーダ (Convolutional Autoencoder:以下 CAE) を用いて画像データに対して次元削減し、低次元に圧縮したデータにクラスタリングを行い、その有効性を検証することを目的とする。

2 畳み込みオートエンコーダ

畳み込みオートエンコーダは、畳み込みニューラルネットワークにおける畳み込み処理、プーリング処理をオートエンコーダに適用したモデルである。エンコーダでは、畳み込みとプーリングを用いてデータの圧縮を行い、デコーダでは畳み込みとアップサンプリングを用いて圧縮したデータを行っている。また、教師なし学習であるため、正解データを用いずに学習が行えるという利点がある。

3 実験方法

本研究では機械学習ライブラリの一つである Pytorch を用いて CAE を実装する。データセットとして Pytorch で提供されている MNIST と、より高次元な Cifar-10 を用いて学習済みモデルを作成し、テストデータを与えてその評価を行う。

次に学習済みモデルのエンコーダで次元削減を行ったデータに対してクラスタリングを行い、その結果をシルエット分析によって評価する。

4 実験結果

まず MNIST に対し、CAE を用いて次元削減を行った。学習回数を 50 回、学習率を 0.5 とした。図 1 に処理を施していない 784 次元の MNIST データ、図 2、図 3 にそれぞれ 392 次元、36 次元に圧縮し、復元した結果を示す。

図 2 より、392 次元への圧縮では元のデータをほぼ再現できており、特徴を捉えられているといえる。また図 3 より、

36 次元への圧縮でも、ぼやけてはいるが認識できている文字も多く大まかに形まで復元できていることから、学習回数や学習率を変えれば認識率は上がると考えられる。

次に Cifar-10 に対し、CAE を用いた。MNIST と同様に学習回数を 50 回、学習率を 0.5 とした。図 4 に処理を施していない 3072 次元の Cifar-10 データ、図 5,6 にそれぞれ 1536 次元、36 次元に圧縮し、復元した結果を示す。

図 5 より、1536 次元への圧縮では色と画像内の物体の形まで特徴を捉えて復元ができている。しかし、図 6 では、図 5 とは異なって白黒画像のようになっており、画像内の物体の形もほぼ捉えられずに復元されていることがわかる。

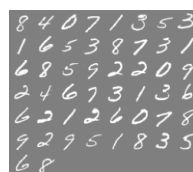


図1 次元圧縮前の MNIST データ



図2 392次元に圧縮後復元した結果



図3 36次元に圧縮後復元した結果

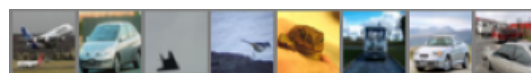


図4 次元圧縮前の Cifar-10 データ

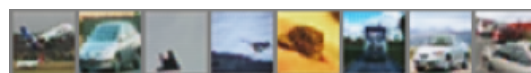


図5 1536次元に圧縮後復元した結果

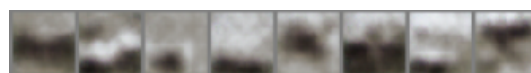


図6 36次元に圧縮後復元した結果

5 おわりに

本研究では、MNIST と Cifar-10 に対し、CAE を用いて学習済みモデルを作成した。

今後は低次元に削減したデータに対して、クラスタリングを行い、その結果を評価する。また、他の次元削減手法についても同様に評価及び比較を行う。

参考文献

- [1] 増井健斗, 金井祐輔, 尾形正泰, 宮崎庸平, 今井倫太, "深層学習の手法を用いたクラスタリング手法の提案," 情報処理学会第 77 回全国大会講演論文集, Vol2015, No.1, pp.347-348, 2015-3-17.