

AutoEncoderによる次元圧縮を用いた高速類似検索手法の提案

安全・安心

情報技術グループ 鈴木 聡
TEL 03-5530-2540

特徴

ニューラルネットワークの一種であるAutoEncoderを用いた次元圧縮によって類似検索を行う技術を開発しました。本技術によって、従来の技術に対して人手を介することなくデータから直接特徴量の抽出が可能になりました。

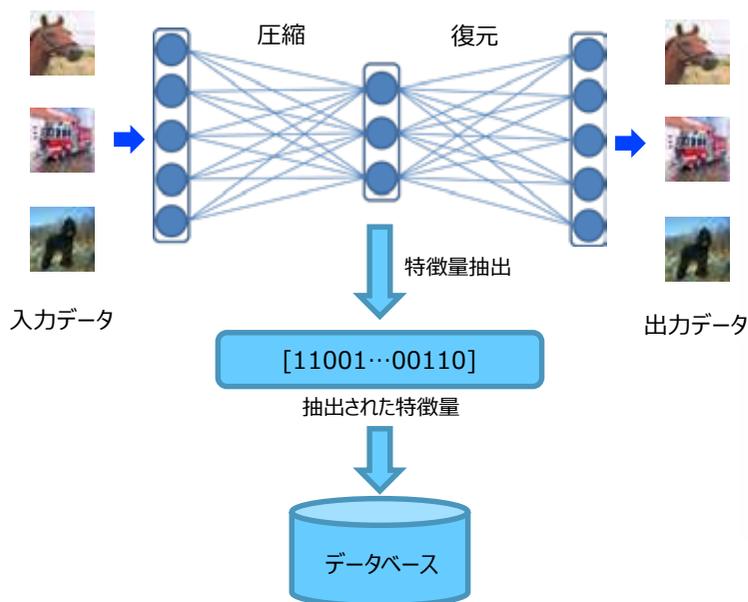


図1. システム概要図

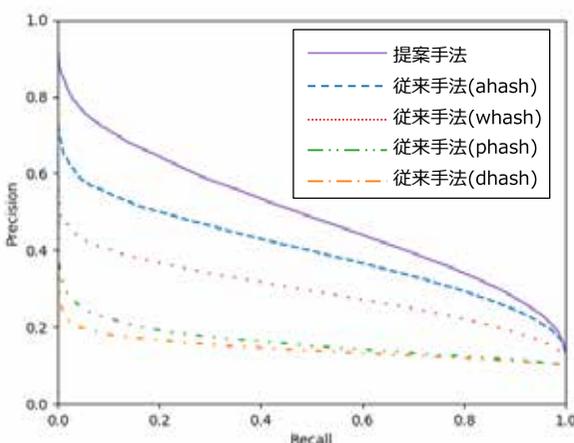


図2. 従来の圧縮技術とのPrecision-Recall曲線での性能比較

AutoEncoderは入力データを圧縮し、圧縮したデータを復元するように学習を行うニューラルネットワークの一種です。本研究では、AutoEncoderを学習した際に圧縮されるデータに着目し、類似検索への応用を目指しました。本技術によって、画像を対象とした場合に、従来の局所特徴と比較して低次元な特徴量を抽出できます。また、データのみを与えることでネットワークが自動的に特徴量を学習するため、人手での特徴設計が不要です。

従来技術に比べての優位性

- 次元圧縮によって人手を介さずにデータの特徴量を抽出できる
- 抽出される特徴量は、従来の手法で得られるものに対して低次元であり、計算コストの削減が可能

今後の展開

- 類似検索アプリケーションへの適用
- 類似度を応用したデータ選別やフィルタリングへの応用

研究者からのひとこと

大量のデータを保有しており、類似検索の利用に興味のある中小企業の方のご相談をお待ちしております。