

複数の非計測センサデータを再現する深層学習モデルの提案

原田 恭兵[†] 長尾 智晴^{††}

[†] 横浜国立大学 理工学部

^{††} 横浜国立大学 大学院環境情報研究院

1. はじめに

工業製品の開発や社内テスト時に、様々なセンサを用いて詳細なデータを取得し、挙動や故障原因などの解析を行う。しかし、高コストであることや搭載スペースの不足により、開発時に使用した、製品への搭載が難しいセンサが多数存在する。製品に搭載可能なセンサのみで、開発時のように詳細な状態がわかれば、高度な解析が可能となり、開発計画への応用が期待できる。複数のセンサデータの再現が可能となれば、製品から取得できるセンサデータで、製品では計測しないセンサデータの再現が期待できる。以上より本研究では、複数のセンサデータを再現するための深層学習モデルを提案する。

2. 提案手法

モデル構造の概要を図1に示す。時系列を加味したセンサデータの特徴抽出および再現を行うために、モデル構造は一次元の Convolutional Neural Network (CNN)と Deconvolutional Neural Network (DCNN)とする。入出力データはある一定の時間間隔 T のセンサデータとする。CNN の入力チャンネル数は入力センサ数 S_{in} とし、複数の入力センサデータを時系列方向に畳み込み、特徴を抽出する。複数のタスクを同時に解くことで、出力同士の共通の要因を学習し、個々のタスクを向上させるマルチタスク学習[1]が効果的だと考え、単一のセンサデータのみでの再現ではなく、複数のセンサデータを同時に再現する。このため、DCNN の出力チャンネル数を出力センサ数 S_{out} とし、CNN で得られた特徴から再現を行う。

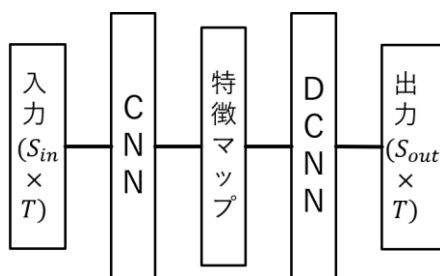


図1. モデル構造

表1. 実験結果(テストデータ)

	提案	Ridge	RF	MLP	BiLSTM
MSE	0.00401	0.01496	0.00520	0.01113	0.01017

表2. 出力センサ数による、ある目的変数の誤差の推移

	提案(1出力)	提案(5出力)	提案(10出力)
MSE	0.0057209	0.0016349	0.0016648

3. 実験設定

本研究では、某企業のセンサデータを使用した。全てのセンサデータは同じタイミングで計測されており、0 から 1 の範囲に正規化されている。具体的なセンサの種類などは秘匿されている。1 データは 200ms にサンプリングされた 40 秒分のセンサデータとし、学習用に 32367 件、テスト用に 8113 件を使用した。

入力データには、20 種のセンサデータを、出力データには再現の対象となる 10 種のセンサデータとした。モデルは入力と同時刻の 40 秒間の出力を行うように学習した。誤差関数には平均二乗誤差(MSE)を、最適化法には Adam を用いた。比較手法には Ridge 回帰、Random Forest (RF), Multi Layer Perceptron (MLP), Bidirectional LSTM (BiLSTM)を用いた。また、複数のセンサを同時に再現する有効性を確認するため、出力センサ数を変更した実験も行った。

4. 実験結果

表 1 に比較結果を示す。提案手法の誤差が、RF, MLP の誤差より小さいことから、時系列を加味した再現が行われたと考えられる。BiLSTM よりも小さいことから、提案モデルの特徴抽出および再現が有効であったと考えられる。

表 2 に提案手法の出力センサ数による、ある1つの出力センサデータの推移を示す。表2より、単一のセンサデータのみを再現するより複数のセンサデータで再現することが効果的であったことがわかる。

5. まとめ

本研究では、非計測マルチセンサデータ再現のための深層学習モデルを提案した。今後は実用時を想定し、さらに多くのセンサを再現できるよう、本手法を改良する。さらに、どのセンサがどのセンサの再現に有用であるかなどの、学習後のネットワークの解析も行う。

参考文献

- [1] LIU, Xiaodong, et al. Representation learning using multi-task deep neural networks for semantic classification and information retrieval. 2015.