

論文の内容の要旨

論文題目 多変量時系列データからの異常検知を目的とした
深層ニューラルネットワークおよび
その学習手法に関する研究

氏名 三木 大輔

1. はじめに

多変量時系列データから異常を検知する技術は、故障診断や状態監視を実現するために重要な技術である。特に深層ニューラルネットワーク (Deep Neural Network, DNN) を用いた手法は、その表現能力の高さから時系列データを解析するうえで有用な手法であるが、実システムに適用するためには事前に膨大なデータを用いた学習によりそのパラメータを最適化することが必要である。そのためには解析する時系列データおよびその異常を決定付ける上で重要な特徴を含む箇所に対してアノテーションを付与し、学習を行うことが必要であったが、データに複雑な特徴が含まれている場合には学習が困難であった。本論文ではこのような課題に対して、マルチインスタンス学習に着想を得た新たに提案する損失関数を最小化するようにDNNの学習を行うことで、時系列データに潜在する特徴を自動的に抽出し、異常を検知可能なDNNを得ることを可能としている。実験では、提案手法により学習されたDNNが異常の検知のみならず定量および識別能力を有することを確認し、さらに本手法を応用することで、軸受の故障原因の識別やその早期検知が可能であることを確認した。さらに映像中の人物動作特徴量を抽出する手法を併せて開発することで人物の行動認識や、通常動作からの逸脱を検知できることを確認した。

2. 多変量時系列解析のための DNN

時系列解析に用いられる一般的な DNN の最適化では、一連の時系列データに対し、その特性を決定付ける上で重要な箇所に対してアノテーションを付与し学習に供することが有効であるが、その作業は煩雑かつ困難なものである。このような課題を解決するため、本研究ではマルチ

インスタンス学習に着想を得たランク学習により DNN の学習を行うため、新たな損失関数を提案する。

2. 1 異常検知型 DNN のための損失関数

図 1 に本手法で提案する DNN の最適化手法の概要を示す。提案する DNN は長さ T の時系列データ $\mathbf{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^T\}$ を入力した際に $\mathbf{y} = \{y^1, \dots, y^T\}$ を出力する構造とし、入力データにおける異常の有無は既知であるが、それらが含まれる箇所は未知であるものとする。DNN の学習では、まず時系列データに異常を含む（正）または、含まない（負）によって各クラスに分割し、それぞれのデータを DNN に入力した際に、正のデータに対して高い値が、負のデータに対して低い値が出力されることを目標に DNN のパラメータを最適化する。各データにおいてその特性を決定付ける上で重要な箇所は未知であるため、ここでは各データにおけるそれぞれ最大の値に着目し、

$$\max_{t \in \{1, \dots, T\}} y_{pos}^t > \max_{t \in \{1, \dots, T\}} y_{neg}^t \quad (1)$$

を満たすように DNN を学習する。ここで、 y_{pos}^t および y_{neg}^t は正および負のそれぞれのデータを DNN に入力した際に出力される $t \in \{1, \dots, T\}$ 時点における推定値である。このような条件を満たす DNN の学習では、以下の損失関数を最小化するように DNN のパラメータを最適化する

$$\mathcal{L} = \max\left(0, 1 - \max_{t \in \{1, \dots, T\}} y_{pos}^t + \max_{t \in \{1, \dots, T\}} y_{neg}^t\right) + \lambda. \quad (2)$$

ここで λ は正則化項であり、DNN の学習で用いるデータへの過学習を防ぎ、学習を安定させるために導入した

$$\lambda = p_1 \sum_{t=1}^{T-1} (y_{pos}^t - y_{pos}^{t+1})^2 + p_2 \sum_{t=1}^T y_{pos}^t. \quad (3)$$

2 つの項はそれぞれ、平滑化項、およびスパース化項である。平滑化項は、時間方向の出力値の変動を低減し、スパース化項は、全体的な出力値を小さくすることで誤検出を低減する効果がある。提案する DNN の学習は(2)式の損失関数の最小化問題となるが、これを解析的に解くことは困難であるため、一般的なニューラルネットワークの学習に用いられる数値的解法により反復的にパラメータを更新することで近似的な解を得る。

2. 2 異常識別型 DNN のための損失関数

先に提案した(2)式を用いた DNN の学習手法を多クラス識別へ応用するため、各時点での DNN の出力を N 次元ベクトルとするほか、損失関数を以下のように一般化する

$$\mathcal{L} = \max\left(0, \sum_{k \in \{k_1, k_2\}} \sum_{n=1}^N (\phi_{kn} - \psi_{kn} Y_{kn})\right) + \lambda. \quad (4)$$

ここで、 $\{k_1, k_2\} \subset \mathcal{K}$ は学習データのインデックスの集合 \mathcal{K} からランダムに選択されたインデックスの対である。 ϕ_{kn} および ψ_{kn} は N クラス識別のために導入された n 次元のラベルであり、以下のように定めた

$$\phi_{kn} = \begin{cases} 1 & \text{if data belongs to } n\text{th class,} \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

$$\psi_{kn} = \begin{cases} 1 & \text{if data belongs to } n\text{th class,} \\ -1 & \text{otherwise.} \end{cases} \quad (6)$$

また、正則化項 λ は

$$\lambda = \lambda_{smooth} + \lambda_{sparse} + \lambda_{entropy} \quad (7)$$

$$\lambda_{smooth} = p_1 \sum_{k=1}^2 \sum_{n=1}^N \sum_{t=1}^{T-1} (y_{kn}^t - y_{kn}^{t+1})^2, \quad (8)$$

$$\lambda_{sparse} = p_2 \sum_{k=1}^2 \sum_{n=1}^N \sum_{t=1}^T y_{kn}^t, \quad (9)$$

$$\lambda_{entropy} = -p_3 \sum_{k=1}^2 \sum_{n=1}^N \phi_{kn} \log \frac{\exp(Y_{kn})}{\sum_m \exp(Y_{km})} \quad (10)$$

であり、1項目、2項目は式(3)と同様にそれぞれ平滑化項、およびスパース化項である。3項目はクロスエントロピー損失項であり、DNNの学習を安定させるために導入した。

4. 実験および考察

本手法により最適化された DNN が異常検知に適用可能であることを検証するため、模擬データおよび実データを用いた検証を行い、異常の検出、定量、および識別能力に関する評価を行った。さらに、回転機器の故障診断および人物行動認識への応用実験を通して実システムへの適用可能性を考察した。

4. 1 実験および考察

提案する DNN およびその学習手法による異常検知が可能か評価するため、まず人工的に生成された波形からの外れ値検知に関する実験を行った。実験では、2つの正弦波の合成波形に対し、平均 $\mu_f = 0$ 、標準偏差 $\sigma_f = 0.05$ のガウス分布に従うノイズを印加した波形（負）と、外れ値としてさらに $\mu_o = 0$ 、 $\sigma_o = 0.20$ のガウス分布に従う外れ値を1%の確率で印加した波形（正）についてそれぞれ10,000組を実験に供した。本実験に用いる DNN は図2に示す3層の DNN とし、隠れ層には128チャンネルを有する一次元 CNN (1DCNN) を用いた。損失関数のハイパパラメータを $p_1 = 10^{-3}$ 、 $p_2 = 10^{-3}$ 、学習係数 10^{-3} とし、Adam optimizer を用い、損失関数(2)を最小化するように10,000回のパラメータ更新を行った。

図3に実験に供した波形および DNN による推定値を示す。外れ値の印加された箇所に対して高い値が推定され、その他の箇所には低い値が推定された。さらに両者の間に相関がみられ、提案する DNN とその学習手法による異常度合の定量可能性が見出された。

4. 2 軸受の故障診断

提案手法による異常の識別能力を評価するため、軸受故障診断のための加速度センサデータを用いた評価を行った。本データセットで用いられた軸受には故障を模擬するため、内輪 (IR)、

転動体 (B) および外輪の 6 時方向、12 時方向および 3 時方向 (OR@6,12,3) に放電加工を施すことで 7 mil の亀裂が加えられている。加速度センサデータは、モーターの両端の 12 時方向の位置に取り付けられた加速度計から取得されている。実験では図 2 の DNN における 1DCNN 層を Long Short-Term Memory (LSTM) 層に代え、より長期的な特徴を抽出できるようにした。損失関数のハイパラメータを $p_1 = 10^{-1}$ 、 $p_2 = 10^{-5}$ 、 $p_3 = 10^{-1}$ とし、学習係数 10^{-3} とし、Adam optimizer を使い、損失関数(4)を最小化するように 10,000 回のパラメータ更新を行った。

図 5 に、それぞれの加速度データおよび推定値を示す。異常に特徴的な箇所に対して高い値が推定され、その他の箇所に対しては低い値が推定された。この結果は提案する DNN および学習方法によって、データに潜在する特徴を自動的に抽出および異常の識別ができることを示している。ここで、提案手法の識別精度を評価するために、長さ T だけのデータから推定される故障クラスを次の式に従って決定し、先行研究と比較した

表 1 に示す通り、他の手法と比較しても良好な識別精度が確認され、本手法が異常の検出および定量的みならず、識別にも適用可能であることが確認された。

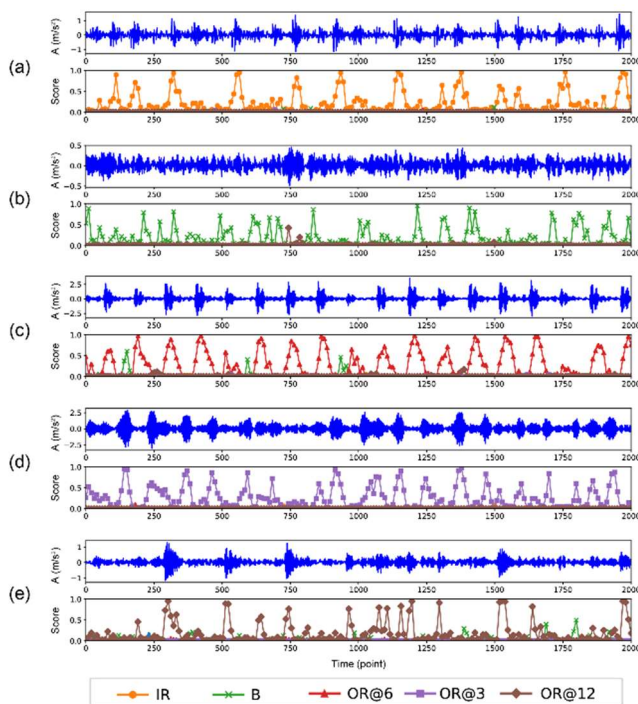


図 5 加速度センサデータおよび DNN の推定値

5. 結論

本研究では、多変量時系列データからの異常検知を目的とした DNN の学習手法を提案し故障診断・映像解析への適用可能性の評価を行った。マルチインスタンス学習に着想を得たランク学習により DNN の学習を行うための新たな損失関数を提案し、アノテーションの付与が困難な多変量時系列データから、データの持つ特性を決定付ける上で重要な特徴を自動的に抽出できることを確認した。さらに映像中の人物動作の特徴抽出手法を併せて開発することで、得られた特徴量から人物動作の通常時からの逸脱を検知することを可能とした。本技術は原子力関連施設における機器の状態監視や映像監視など、様々な分野への応用が可能と考えられる。