

深層学習に基づく生産設備の異常診断技術

千 田 有 一*



1. はじめに

生産設備の予期しない故障は生産計画に大きな影響を与えることから、装置の異常診断や予測技術の開発が期待されています。一方、近年、深層学習（ディープラーニング, DL）の関連技術が急速に発展し、生産装置の異常診断・予測への応用も期待されています¹⁾。これらの技術では、診断モデルの構築に活用するデータの質と量が非常に重要です。つまり、正常時と異常時でのデータを十分に取得でき、さらにその差異の特徴が明確である場合には異常判別が容易になります。しかしながら、多くの場合には異常は減多に起こらず、異常時のデータの確保が困難である場合が多くなります。このことが、装置の異常診断・予測を難しくしています。これに対処するためには、正常データの確率分布を参考にして異常時データを人工的に生成するアプローチ、あるいは、自己符号化器（オートエンコーダー）²⁾などの教師なし学習のアプローチを取るなどの方法が必要になります。これに対して、筆者らの研究グループでも、上記の方法について検討しています^{3,4)}。本稿では、これらの方法の適用例および、生産現場での適用例についてご紹介します。

2. 異常データの人工的生成とマルチラベルディープニューラルネットワークによる異常検知³⁾

ここでは、部品洗浄装置における処理プロセス異常の検知について紹介します。その部品洗浄プロセスでは、複数の処理プロセスを経て部品洗浄を行います。今回は、各プロセスにおける作業時間の変化に着目して異常の検知可能性を検討しました。すなわち、一つの作業プロセスに要する時間が標準時間から乖離している場合には何らかの問題を内包しているとしてそれをいち早く検知しようというものです。その場合、一つ一つのプロセスが完全に独立である場合には、それぞれのプロセス処理時間に着目

すればよいこととなりますが、関連する処理プロセスに影響を及ぼす場合には、それぞれのプロセスを独立に考えるよりも、相互の関連性を考慮できる構造の方がより適切に異常検知可能と考えられます。そこで、マルチラベルディープニューラルネットワーク（ML-DNN）を適用しました³⁾。

ML-DNNは、ディープニューラルネットワーク（DNN）の複数の出力層それぞれを二値分類する構成方法です。設定したML-DNNでは、部品洗浄プロセスを構成する複数の処理プロセスにおける各処理時間を入力とし、各プロセスでの異常可能性を正常か異常かの二値の出力で表します。その出力を、構成される複数の処理プロセスすべての数の出力について考えた場合には、複数出力における二値分類の問題となり、ML-DNNが適用できます。それにより、洗浄工程における処理プロセス間の影響を考慮しつつ、どのプロセスで異常が発生しているかを検知しようと考えました。

その場合、十分な量の正常データと異常データによってML-DNNを学習させる必要が有ります。このとき、正常データは通常の稼働時のデータが利用可能であるのに対し、異常時のデータは入手が難しい現実が有ります。この問題を回避するため、正常時のデータの分布を求めておき、その分布から外れた値を異常データとして人工的に設定し、それらのデータに基づいてML-DNNの学習を行うアプローチとしました。

例えば、洗浄工程内の処理プロセスに要した時間のヒストグラムから確率分布を調べた結果の例を図1に示します。確率分布としてはさほど単純なものではないことが分かり、確率分布モデルの構築が必要となります。しかしながら、今回は最も簡便な方法として正規分布をあてはめることとしました。まず、正規分布であると仮定して平均 μ と標準偏差 σ を求めます。もし、正規分布であれば、 $\mu \pm 3\sigma$ の外側のデータの発生確率は0.3%程度であることから、 $\mu \pm 3\sigma$ 以内が正常値であり、 $\mu \pm 3\sigma$ から $\mu \pm 8\sigma$

*信州大学 学術研究院（工学系）教授

の範囲のデータが異常値であると仮定し、その確率分布によって異常データを人工的に生成しました。その上で、ML-DNNの学習用データとして用いました。その際、正常データとしては通常の稼働時のデータを用いました。

以上の方法で生成したデータを用いてML-DNNを学習させ、異常検知システムを構築しました。その結果、異常判別についての正答率99.6%を実現できることを確認しました。一方、通常のDNNを用いた場合の正答率は98.2%であったことから、ML-DNNによる方法の優位性が確認されました³⁾。今後は、確率分布モデルとして混合ガウスモデルなどの適用を検討しています。

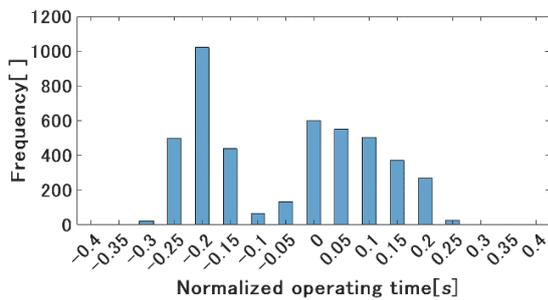


図1 処理プロセス時間のヒストグラム³⁾

3. スペクトル特性と自己符号化器を併用した異常検知⁴⁾

ここでは、部品洗浄装置における高圧ポンプの異常検知の検討例を紹介します⁴⁾。ポンプの圧力データは、その駆動周波数の影響が大きいと考えられることから、フーリエ変換によって得られた周波数スペクトルに基づいたアプローチが有効であると考えられます。そこで、時系列データをFFTすることによって周波数特性を求め、その周波数スペクトルの変化による異常検知を目指しました。ただし、この場合でも、異常状態でのデータ取得が容易ではないことが問題となります。そこで、教師なし学習の一つである自己符号化器³⁾に基づいた方法を用いました。

自己符号化器では、異常データは必ずしも必要では無く、正常データのみを用いて学習させて異常判別器を構築することができます。すなわち、図2のようなネットワーク構造を設定し、符号化 (encode) と復号化 (decode) を用いて入力を再現する \hat{x} を出力させます。その際、入力と出力のデータ値の距離を損失関数として設定し、その値が最小となるように学習させます。これにより、適切に学習させることができた場合には、正常データを入力させた場合、損失関数値が有る程度小さい値となります。そこで、

その値を目安として閾値を設定します。すると、入力したデータによって得られる損失関数値が閾値より小さければ正常状態であり、大きければ異常状態であると判別できます。以上の考え方によって自己符号化器を構成し、同時に閾値を設定して異常判別器を構成しました。

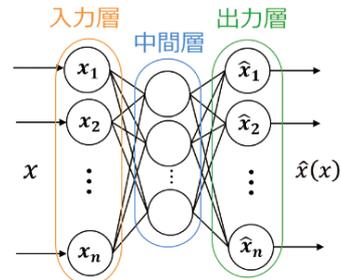


図2 自己符号化器²⁾

設定した自己符号化器の入力はFFTによって求めた127点の周波数スペクトルデータとしています。そして、自己符号化器の入力と出力の差の2乗和を損失関数として設定しました。一方、評価のためには異常時の周波数スペクトルデータが必要になりますが、それは正常データの一部を加工することによって人工的に作成しました。すなわち、正常時の周波数スペクトルデータのうち、一部の周波数スペクトル値を増減させて異常と想定したデータを作成しました。

以上の手順で自己符号化器を構成し、性能を検証した結果を図3に示します。図3において、灰色(右目盛)は周波数スペクトルを表し、赤線(左目盛)は異常検知率を示します。横軸は周波数をスケール変換して示しています。図3の周波数幅全体での異常検知率は約83%であり、概ね検知できているようにも見えます。しかし、図3における約0.8の周波数において異常検知率が局所的に悪くなっています。同様の周波数は約1.5, 1.8, 2.3の周波数近辺にも見られます。これらの周波数では、正常データにおいて周波数スペクトルの振幅が大きくなっており、その周波数では振幅の変化の検知精度が悪くなってしまうと考えられます。

この原因としては、正常データにおいても、振幅の大きな周波数において、図4のパターン1に示すように周波数値が一つに集約されるわけではなく、パターン2や3のように、その前後にばらついてしまうため、それらの周波数帯域でのスペクトルの振幅値変化の検出が難しくなってしまったと考えられます。この点を改善させるため、図4に示す周波数スペクトルのパターンを同一のパターンであるとし

て取り扱うことが適切と考えられます。そのため、異常検知率の低かった特定の周波数帯域については、自己符号化器による判別に加えて追加の判別機能を付加することを考えました。

そこで、図5に示す判別フローに変更しました。図5では、自己符号化器によって正常と判別された場合には、検知精度の悪い周波数帯域に異常の可能性がどうか追加で判別しています。異常検知精度の悪い周波数帯域では、図4のパターン2や3であっても、近傍のすべてのピーク値が含まれる程度の周波数帯域幅を設定し、その帯域幅でのパワースペクトルの和に基づいて異常の有無を確認しました。その手順としては、まず、設定した帯域幅によるパワースペクトル和を正常データについて求めます。次にそのヒストグラムによってパワースペクトル値の分布を求め、正規分布と仮定して平均値 μ と標準偏差 σ を求めます。その上で、第2章の考え方と同様に、データ値が $\mu \pm 3\sigma$ 以内が正常値と考えて閾値を設定し、正常か異常かの判別を行うこととしました。この方法は、自己符号化器で誤判別の多かった周波数帯域（図3では約0.8, 1.5, 1.8, 2.3などの周波数帯域）でそれぞれ用いました。つまり、まず自己符号化器による判別で正常であると判別された場合、周波数0.8近傍の帯域での判別を行います。そして、正常と判別された場合には、さらに周波数1.5近傍の帯域での判別を行うなど周波数を順次変更して判別しました。以上の方法を用いた場合の性能検証結果を図6に示します。図3と同様のデータにおいて、周波数幅全体での異常検知率は図3の結果である83%から約92%に向上させることができました。さらに、図3において異常検知率の低かった周波数帯域での検知精度が向上していることが確認できます。

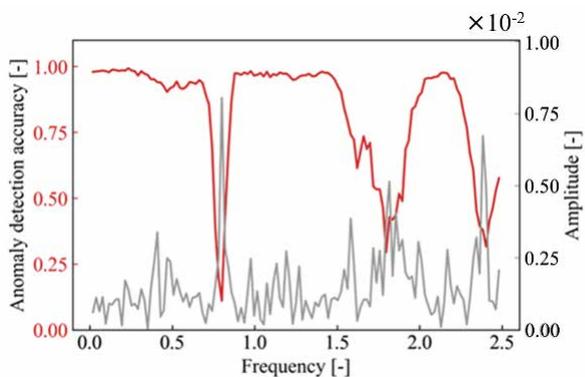


図3 周波数データと異常検知率 (改善前)⁴⁾

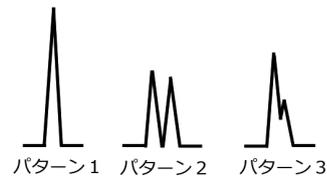


図4 周波数ピーク値の典型的なパターン例⁴⁾

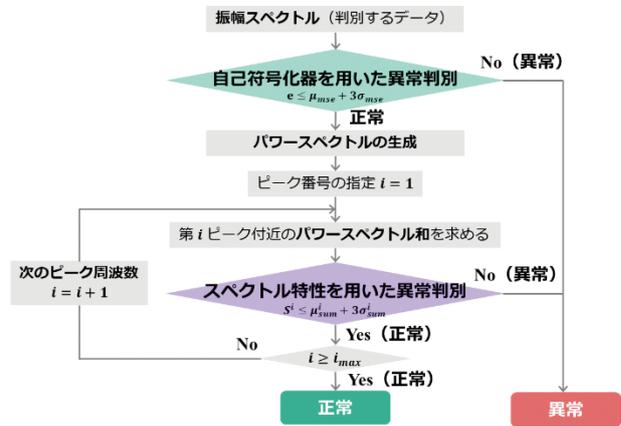


図5 異常検知手順⁴⁾

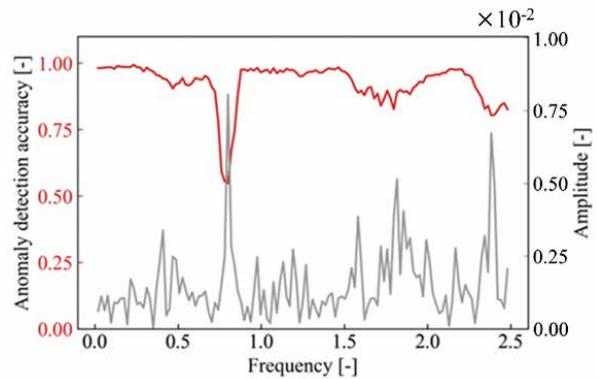


図6 周波数データと異常検知率 (改善後)⁴⁾

4. 生産現場での応用例

第2章で述べた、ML-DNNを用いた異常検知システムが生産現場設備の実運用で有効かの確認を実施しました。検証は、カヤバ株式会社デジタル変革推進本部システム開発室に実施いただきました。確認方法は、現場が管理しやすい動作時間に閾値を設け、閾値を超える事象が発生するか観測を試みました。閾値越えが発生した際、ML-DNNを実施し同等の異常結果となれば実運用時における異常検知にも対応可能と考えられます。

閾値越えしたデータ件数は一目でわかる様、図7のようにBIツールであるTableauを活用して見える化を実施し、定期的にデータ状態を確認しました。すると、ある期間において、「ワーククランプ横解除」動作で異常件数が増加する現象を2度捉えることができました。異常は同期間で2回発生し、図8の管

理図で確認しても異常は顕著に表れています。同じデータを用いて、図9のようにML-DNNを実施し管理図と比較しました。その結果、異常傾向が一致することが確認できました。

このデータから実設備を確認しました。部位は「ワーククランプ横解除」となるため、関連した場所に異常が無いか確認したところ、異常1はエアチューブからのエア漏れ(写真1)、異常2はエアチューブ劣化(写真2)が確認され、交換を実施しました。交換後は異常データが解消されることも確認できました。

高圧ポンプの圧力データにおけるスペクトル特性と自己符号化器を併用した異常検知については、24年度以降に実用化に向け準備すると共に、振動、音データでも有効性が確認できるかを継続して検証していく予定です。

5. まとめ

本稿では、DL等のデータに基づいた生産装置の異常診断・予測技術について、マルチラベルディープニューラルネットワークや自己符号化器を用いた方法を紹介しました。これらの技術は、異常データの入手が難しい状況でも適用できます。その際、異常検知の際に着目すべき特徴量の抽出のために用いる信号処理技術なども重要な要素となっています。さらに、装置の異常要因の考察やその影響がどのような形で現れるかについての考察も重要です。その意味では、対象とする装置のモデルを構築し、そのモデルからの乖離度合いによって異常診断や予測を行っていることと等価となります。この技術は、制御工学の分野ではシステム同定⁵⁾とよばれており、動的システムの異常診断技術構築には、システム同定技術の知見も応用可能であると考えられます。

参考文献

- 1) 日本機械学会, 第21回評価・診断に関するシンポジウム予稿集 (2023)
- 2) 岡谷貴之: 深層学習 (改訂第2版), 講談社 (2022)
- 3) 西田, ほか: マルチラベルディープニューラルネットワークを用いた機械部品製造装置の異常検知, 日本機械学会D&D2021 (2021)
- 4) 坂, ほか: スペクトル特性と自己符号化器を併用した生産設備の異常検知, 日本機械学会D&D2023 (2023)
- 5) 足立修一: システム同定の基礎, 東京電機他大学出版社 (2009)



図7 閾値越えした異常時間の件数

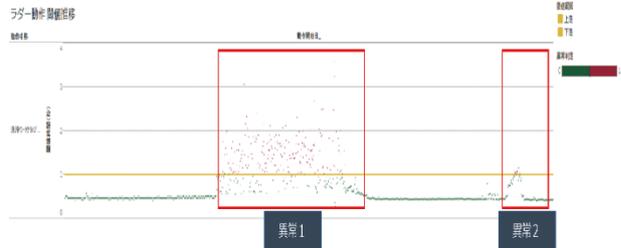


図8 管理図を用いた異常検知グラフ

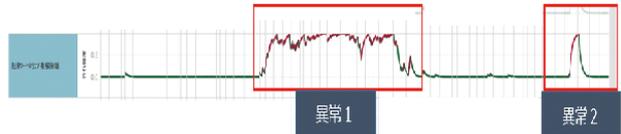


図9 ML-DNNを用いた異常検知グラフ



写真1 異常1のエア漏れ箇所



写真2 異常2のチューブ劣化箇所