

機械製品の動作音のエネルギー変動周期性に基づく 異常検知手法

松井 威¹ 石田 繁巳^{2,a)} 金井 貴浩³

概要：

機械製品の品質検査において、動作音による異常検知の自動化が強く求められている。本稿では、回転機器などの動作音に見られるエネルギー変動の周期性に着目し、音響信号から周期特徴を抽出する特徴量と、オートエンコーダを用いた教師なし異常検知手法を提案する。本手法は、機械動作の本質的なリズムからの逸脱を直接的に評価するアプローチをとる。回転・往復機構に現れる衝撃音の周期性を表す特徴量を抽出し、製品の動作状態ごとに正常時の特徴を学習したオートエンコーダモデルを構築する。そのモデルの再構成誤差を用いることで異常検知を行う。実データを用いた評価の結果、ROC AUC スコアは 0.599 となった。詳細分析の結果、異常データはその検出難易度により二極化することが確認された。想定した音響的特徴を持つ異常データに対しては、正常データ群から乖離した高い Z スコアを示し、安定して検知可能であることを実証した。一方で、一部の異常サンプルについては正常サンプルとの区別が困難であることも明らかとなった。以上の結果より、本手法は一部の音響的特徴を伴う異常に対しては有効性を有していることを確認した。

キーワード：異常検知, オートエンコーダ, 音声信号

1. はじめに

機械製品の製造現場において、製品の品質を担保するための動作音検査は極めて重要な工程である。音響信号は製品内部の微細な構造的欠陥や組立不良を非接触かつ高感度に反映するため、従来より熟練検査員による官能検査が広く行われてきた。しかし、官能検査は検査員の体調や経験に依存する属人性の問題や、検査コストの増大、判定基準の定量化が困難であるといった課題を抱えている。

これに対し、近年の深層学習技術の発展に伴い、正常データのみを学習して異常を検知する教師なし異常検知技術の導入が進められている。一般的には、音響信号をスペクトログラムなどの画像特徴量に変換し、汎用的な画像認識モデルを用いて正常データの特徴を学習させるデータ駆動型のアプローチが主流である。しかし、こうしたアプローチは、高い表現力を持つ反面、モデルが大規模化しやすい。

一方で、従来の信号処理ベースの手法では、対象製品の物理特性に基づいた特徴抽出が行われてきた。機械製品の異常発生メカニズムに立ち返ると、Tang ら [1] や Wodecki

ら [2] が指摘するように、モーター、ギア、ベアリングといった回転・往復機構を有する多くの機械製品において、部品の接触や欠損に起因する異常は、「短い衝撃音が特定の周期で連続して発生する」という共通した物理的特性を有している。これらの信号処理手法では、この周期性や衝撃成分に対して人間が閾値を設定していたが、このアプローチは個体差や環境ノイズの影響を受けやすい。また、物理特徴を特定の周波数帯の面積や変動幅の最大値といった単一の指標に圧縮し、それに対して人間が閾値を設定して判定することが一般的であったが、このように情報を単純化するアプローチでは、正常と異常を区別するための形状情報が欠落しやすい。すなわち、単一の指標に圧縮された特徴量に対して閾値を設定する従来のアプローチは、個体差や環境ノイズの影響を排除しきれず、製品ごとに厳密なチューニングを要するという課題があった。

したがって、こうした個体差や環境ノイズによるゆらぎを許容し、厳密なチューニングなしに異常を検知するためには、物理的特性を単なる数値として圧縮するのではなく、正常なリズムの分布形状として保持したまま評価する必要がある。そこで本稿では、「エネルギー変動の周期性」を多次元のスペクトル形状として捉える特徴量を新たに設計し、その正常な変動パターンの学習と異常度の算出をオー

¹ 公立はこだて未来大学院 システム情報科学研究科

² 公立はこだて未来大学 システム情報科学部

³ アイシン・ソフトウェア (株)

a) ish@fun.ac.jp

トエンコーダに委ねる手法を提案する。異常の物理的特性が現れる特徴量をパターンとして入力することで情報の欠落を防ぎ、かつ判定基準をデータから学習させることで、個体差や環境変動に対しても柔軟で自動的な検知の実現をする。

提案手法の有効性を検証するため、実際の機械製品の動作音データを用いた評価実験を行った結果、ROC AUC スコアは 0.599 となった。詳細な分析により、異常データはその検出難易度によって明確に二極化することが確認された。想定した音響的特徴を伴う異常を持つ異常データに対しては、正常範囲から有意に乖離した高い異常スコアを示し、提案手法が安定して検知可能であることを実証した。一方で、一部のデータについては正常データの変動範囲内に異常スコアが留まり、今回着目した物理特徴量では正常動作との区別が困難であることも明らかになった。本研究の結果は、物理モデルに基づくアプローチが特定の異常サンプルに対しては有効であるとともに、一部の異常サンプルの異常特徴が本手法で抽出した特徴量では捉えられておらず、正常サンプルとの区別が困難であったことを示唆している。

本稿の構成は以下の通りである。2章では、音響異常検知に関する関連研究とその位置付けについて述べる。3章では、提案するエネルギー変動スペクトルと単純化オートエンコーダを用いた異常検知手法の詳細を説明する。3章では、実機データを用いた評価実験を行い、提案手法の有効性と検出結果に基づくデータセットの分析について論じる。最後に5章で本稿をまとめる。

2. 関連研究

機械製品の動作音検査における異常検知技術は、製造現場の自動化や効率化の観点から重要な課題である [3]。従来、この分野では熟練検査員の聴覚に頼る官能検査が主であったが、定量化と自動化を目指して様々な信号処理手法が検討されてきた。初期の取り組みとして、大瀧らはプラスチックギアの異音検査において、FFTを用いた周波数解析や波形の実効値 (RMS) 監視に基づく定量的な判定手順を開発した [4]。また、笠原らは、動作音の周波数が変動するような非定常な異音に対して、ケプストラム解析と周波数リッジ解析を組み合わせることで、変動する周波数成分をロバストに追跡・抽出する手法を開発している [5]。このような物理的な特性に基づいた特徴量設計は、異常音を本質的に区別する上で依然として重要である。本研究も物理特性に基づくアプローチをとるが、従来の信号処理手法が物理的特徴を特定の周波数成分や統計量といった単一の指標に圧縮し、閾値判定を行っていたのに対し、本手法では機械特有の周期性をスペクトル形状として捉え、その変動パターン全体を機械学習モデルに学習させる点が異なる。

近年、深層学習技術の発展に伴い、正常データのみを用

いてモデルを学習させる教師なし学習アプローチが主流となっている。中でも最も代表的なのがオートエンコーダを用いた手法であり、Duman らは産業プロセスにおける音響データに対して畳み込みオートエンコーダを適用し、その有効性を実証している [6]。オートエンコーダベースの手法は、正常音を再構成するように学習し、異常音入力時の再構成誤差を異常スコアとするものであるが、異常音に対しても誤差が小さくなるリスクがある。この課題に対し、Koizumi らは統計的仮説検定のネイマン・ピアソンの補題に基づき、低い誤検知率 (FPR) 制約下で検出率 (TPR) を最大化する目的関数を提案し、オートエンコーダの学習を最適化している [7]。また、音響信号の時系列的な特徴を捉えるために、Marchi らは双方向 LSTM を組み込んだデノイズングオートエンコーダを提案し、高い検出性能を達成している [8]。

一方で、再構成誤差に依存しないアプローチも活発に研究されている。Primus らは、入手が容易な無関係の音データを「擬似的な異常 (Proxy Outliers)」として利用し、正常音とそれ以外を分類する二値分類問題として定式化する手法を提案した [9]。Kuroyanagi らはこの考えを発展させ、二値分類モデルにクラス重心からの距離学習 (Metric Learning) を導入することで、特徴空間における正常クラスの凝集度を高め、識別性能を向上させている [10]。また、松本らは多クラス分類器と中間層出力の再構成器を組み合わせることで、分類と再構成の両方の利点を統合した手法を検討している [11]。さらに、生成モデルを用いたアプローチとして、Haunschmid らは正規化流 (Masked Autoregressive Flow: MAF) を用いた密度推定手法を提案しており、機械の種類に応じて異常スコアの集約方法を使い分けることの重要性を示している [12]。

これらの一般的な深層学習アプローチの多くは、高い表現力を持つ反面、モデルが大規模化しやすい。また、音響信号をスペクトログラムなどの「画像」として処理するため、微細なテクスチャの違いを捉えることには長けているものの、機械特有の「動作サイクルと連動した衝撃音の周期性」という時間的な構造情報が、画像特徴の中に埋没してしまうという懸念がある。

本稿では、このような背景を踏まえ、機械の回転や往復運動のリズムに合わせて繰り返し発生する衝撃音に基づく特徴量と教師なし学習を組み合わせることで、異常の発生メカニズムを的確に捉えつつ検知の自動化を図る。周期成分の分析に特化したエネルギー変動スペクトルとオートエンコーダの再構成誤差を用いた手法の有効性を検証する。

3. エネルギー変動周期性に基づく異常検知手法

3.1 アプローチ

本手法のキーマイディアは、連続する短い衝撃音の周期性

に着目することである。機械内部の部品同士の接触や欠損などに起因する異音の多くは、短い衝撃音が一定の間隔で連続して発生している状態である。この衝撃音は特定の音程を持たない広帯域なノイズであり、ログメルスペクトログラム上では、全周波数帯域に広がる「縦方向の線」として現れる。それが連続することで縦縞模様のようなパターンを形成する特徴がある。

本手法では、連続する衝撃音によって生じるログメルスペクトログラム上の縦縞模様を検出する特徴量として、エネルギー変動スペクトルを定義する。エネルギー変動スペクトルは、ログメルスペクトログラムの各時刻で全周波数帯域のパワーを合計した「エネルギー時系列」に対して周波数分析した結果である。エネルギー時系列はログメルスペクトログラムの周波数軸の情報を捨てた「その瞬間の音のエネルギーの大きさ」であり、ログメルスペクトログラム上の縦縞模様はエネルギー時系列ではパルス列として現れる。そこで、エネルギー時系列の周波数成分をFFTにより分析し、ログメルスペクトログラム上の縦縞模様の周期性を取り出す。

機械に異常がある場合、エネルギー変動スペクトルには正常な機械で見られない周期性が生じる。正常な機械であっても、モータの回転に起因する定常的な動作リズムが存在するため、エネルギー変動スペクトルには特定の周波数成分が確認される。機械の異常によって生じる連続した衝撃音は、正常な機械で見られる周期性とは異なる周期性としてエネルギー変動スペクトル上に現れる。そこで、正常な機械のエネルギー変動スペクトルをオートエンコーダに学習させ、正常なスペクトルからの逸脱をオートエンコーダの再構成誤差として取り出すことで、異常を検知する。

エネルギー変動スペクトルは短時間フーリエ変換によって得られるスペクトログラムからも計算できるが、本研究ではメル尺度に変換してからエネルギー変動スペクトルを算出した。メル尺度に変換してからエネルギー変動スペクトルを算出することで、高周波帯域のノイズ成分を抑制しつつ、官能検査の基準となる人間の聴覚特性に合わせたエネルギー変動を抽出できる。

3.2 概要

図 1 に、エネルギー変動周期性に基づく異常検知手法の概要を示す。本手法は、音響信号から衝撃音の周期性を抽出する特徴量抽出部分と、その特徴量の正常パターンをオートエンコーダモデルの構築して学習する部分、オートエンコーダの再構成誤差を用いた判定部分から構成される。各部分は以下の5つのステップに分けられる。

- 特徴量抽出:
 - (1) ログメルスペクトログラムの抽出
 - (2) エネルギー時系列の抽出

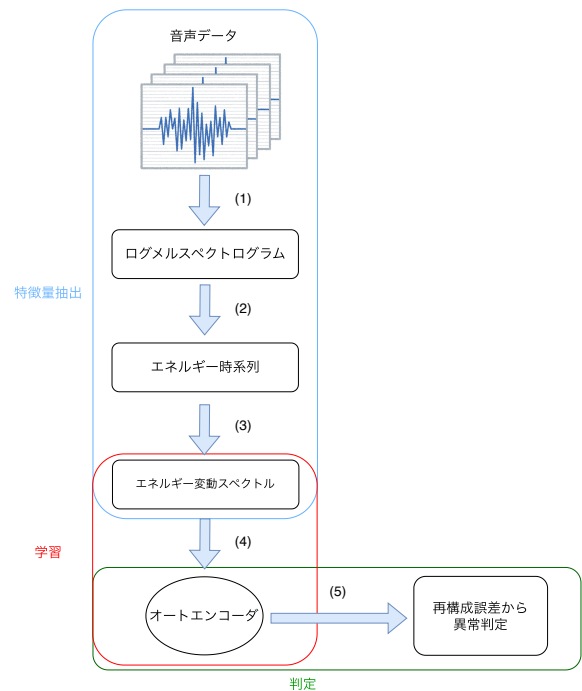


図 1 エネルギー変動周期性に基づく異常検知手法の概要

- (3) エネルギー変動スペクトルの抽出
- 学習:
 - (4) オートエンコーダの学習
- 判定:
 - (5) 再構成誤差による異常判定

まず、特徴量抽出部分であるステップ (1) から (3) では、正常および検査対象の音声データから特徴量としてエネルギー変動スペクトルを抽出する。次に、学習部分であるステップ (4) では、正常データから抽出されたエネルギー変動スペクトルを用いてオートエンコーダの学習を行う。最後に、判定部分であるステップ (5) において、検査対象データの特徴量を学習済みのオートエンコーダに入力し、入力スペクトルと復元スペクトルとの差分を再構成誤差として計算する。本手法では、この再構成誤差を異常スコアとして定義し、正常・異常の判定を行う。

以降では各ステップについて詳述する。

3.3 ログメルスペクトログラムの抽出

本ステップでは、入力された音声波形データをウィンドウに区切り、短時間フーリエ変換およびメルフィルタバンクを適用して振幅を対数変換することでログメルスペクトログラムを抽出する。周波数分解能に対応するメル周波数ビン数を F 、短時間フーリエ変換を適用するウィンドウの数を T とすると、ログメルスペクトログラムは F 行 T 列の行列である。この行列を $X = [x_{ij}]$ とすると、要素 x_{ij} ($i = 1, \dots, T, j = 1, \dots, F$) はウィンドウ i における

j 番目のメル周波数の音響エネルギーの強さを表している。

3.4 エネルギー時系列の抽出

本ステップでは、3.3 で得たログメルスペクトログラムから周波数情報を圧縮してエネルギー時系列へと変換する。エネルギー時系列は、各ウィンドウにおける音のエネルギーの総量を表すものである。エネルギー時系列ではログメルスペクトログラム上で縦に広がっていた衝撃音の様子が鋭いパルス波形として単純化され、後段の周期性解析に適した形式となる。

エネルギー時系列への変換は、ログメルスペクトログラムの行列 X の列ごとの総和をとることで行われる。すなわち、エネルギー時系列 $e = [e_1, e_2, \dots, e_T]$ の要素 e_i は以下のように計算される。

$$e_i = \sum_{j=1}^F x_{ij} \quad (i = 1, \dots, T) \quad (1)$$

3.5 エネルギー変動スペクトルの抽出

本ステップでは、3.4 で得られたエネルギー時系列 $e = [e_1, e_2, \dots, e_T]$ をフーリエ変換し、エネルギー変動スペクトルを抽出する。エネルギー変動スペクトルはエネルギー時系列の周波数成分を表す特徴量である。エネルギー時系列 e は音のエネルギーの時間変化を表していることから、これをフーリエ変換することでエネルギー変動スペクトルが得られる。エネルギー変動スペクトルは音のスペクトルを表すものではなく、音のエネルギーの強弱のリズムが1秒間に何回繰り返されるかという周波数を表しており、そのピークは機械の回転や往復運動のサイクルに対応する。

エネルギー変動スペクトルの抽出では、エネルギー時系列の直流成分を除去、すなわち平均値を差し引いた上で窓関数としてハンニング窓を適用し、高速フーリエ変換を適用する。長さ T のハンニング窓を $w = [w_1, \dots, w_T]$ とすると、直流成分を除去してハンニング窓を適用したエネルギー時系列 $e' = [e'_1, \dots, e'_T]$ の要素 e'_i ($i = 1, \dots, T$) は

$$e'_i = w_i(e_i - \bar{e}) \quad \left(\bar{e} = \frac{1}{T} \sum_{i=1}^T e_i \right) \quad (2)$$

となる。 e' を高速フーリエ変換してエネルギー変動スペクトル $S = [s_1, \dots, s_{T/2}]$ を得る。

3.6 オートエンコーダの学習

本ステップでは、正常な機械のエネルギー変動スペクトルが示すパターンをオートエンコーダで学習する。オートエンコーダは、入力された特徴量ベクトルを忠実に復元するように設計されたニューラルネットワークである。本ステップでは、正常なデータから抽出したエネルギー変動スペクトルのみを用いてオートエンコーダを学習する。

学習を通じて、オートエンコーダは正常データが持つスペクトル形状の構造を効率的に圧縮・復元する能力を獲得する。これにより、オートエンコーダは正常なデータに対しては小さな復元誤差を出力するようになる。学習データには含まれない異常データ特有の異常なパターンのエネルギー変動スペクトルが入力された場合、オートエンコーダはその異常なパターンを正しく復元できず、入力と出力の間に大きな誤差を生じさせる。

3.7 再構成誤差による異常判定

本ステップでは、オートエンコーダの再構成誤差に基づいて異常スコアを算出し、異常スコアがあらかじめ設定した閾値より高い場合に異常と判定する。検査対象データのエネルギー変動スペクトルをオートエンコーダに入力し、再構成されたスペクトルを出力する。そして、再構成されたスペクトルと入力したエネルギー変動スペクトルの再構成誤差、すなわち平均二乗誤差を計算して、異常スコアとする。

異常判定の閾値は、オートエンコーダの学習における検証データに対する異常スコアの分布から定める。学習済みのオートエンコーダに対して学習で使用した検証データを入力し、異常スコアを得る。検証データは正常な機械であることから、異常スコアは十分に小さくになると考えられる。得られた異常スコアの分布に対して十分に大きい閾値を設定することで異常判定を行う。閾値の設定は異常検知性能に大きな影響を与える。このため、本稿では閾値を変化させたときの ROC-AUC を用いて、提案手法の性能を検証する。

4. 評価

4.1 データセット

評価には、本研究で収集した、製品の動作音データを用いた。対象とする製品は往復動作を行う機械であり、各サンプルは1台の製品の1回の往復動作に対応して取得している。往復動作は OPEN と CLOSE の2つの動作状態から構成され、それぞれについて音声データを収集し、同じサンプルの音声データは同じ番号となるように各データに連番をつけた。

データは正常データと異常データに分けて収集しており、ギアの歯打ち音の間隔に異常が現れるような、周期的な異常音を用意した。

収集した音声データは、サンプリング周波数 51.2 kHz、量子化ビット数 24 bit でデジタル化されている。収集された全てのデータは 5.0 秒以上の長さを有していたため、前処理として、一律に先頭から 5.0 秒間 (256,000 サンプル) を切り出すことで、全てのサンプルのデータ長を統一した。以上の収集および前処理を経て、本実験で用いる評価用データセットを作成した。

表 1 データセットの内訳

ラベル	OPEN	CLOSE	計
正常	1000	1000	2000
異常	6	6	12

作成したデータセットに含まれる正常データおよび異常データのサンプル数は表 1 に示す。

4.2 評価環境

作成したデータセットは、以下のように学習データ、検証データ、評価データに分割した。評価データには収集された異常データの全数である 6 データに対し、同数の正常データ 6 データを割り当て、計 12 データ（正常: 6 サンプル、異常: 6 サンプル）を使用した。学習時のモデル性能評価および学習の早期終了（Early Stopping）の判定基準として使用する検証データには、評価データを除いた正常データの 10% を割り当てた。学習データには、これら評価データおよび検証データを除いた、残りのすべての正常データを使用した。

対象製品の動作音は OPEN 動作および CLOSE 動作の 2 つの状態において音響的特徴が異なる。そのため、両動作を同一のオートエンコーダモデルで学習させるのではなく、OPEN 用と CLOSE 用でそれぞれ個別のモデルを構築し、独立して学習を行った。学習時の損失関数には平均二乗誤差を使用し、最適化アルゴリズムには Adam を採用した。また、モデルの過学習を抑制し、汎化性能を確保するため、Early Stopping を導入した。具体的には、エポックごとに検証データの損失を監視し、10 エポック連続で改善が見られない場合に学習を停止し、検証損失が最小となった時点のモデル重みを採用した。

各オートエンコーダモデルから出力される再構成誤差のスケールや分布は、OPEN 用モデルと CLOSE 用モデルでそれぞれ異なる。そのため、再構成誤差をそのまま用いた場合、動作状態によって異常の基準が異なり、統一的な比較や閾値設定が困難となる。そこで本研究では、異常度として Z スコア（標準化スコア）を採用した。学習データにおける再構成誤差の平均を μ 、標準偏差を σ としたとき、入力データに対するモデルの出力再構成誤差 x の異常度 $A(x)$ は以下の式 (3) で定義される。

$$A(x) = \frac{x - \mu}{\sigma} \quad (3)$$

このように各モデルの学習データ統計量を用いて正規化を行うことで、動作の種類（OPEN/CLOSE）に依存せず、統一的な指標での異常判定を可能とした。

実際の製品検査においては、OPEN 動作と CLOSE 動作のいずれか一方でも異常が生じていれば、その製品は不良品とみなされる。したがって、製品単位での最終的な異常

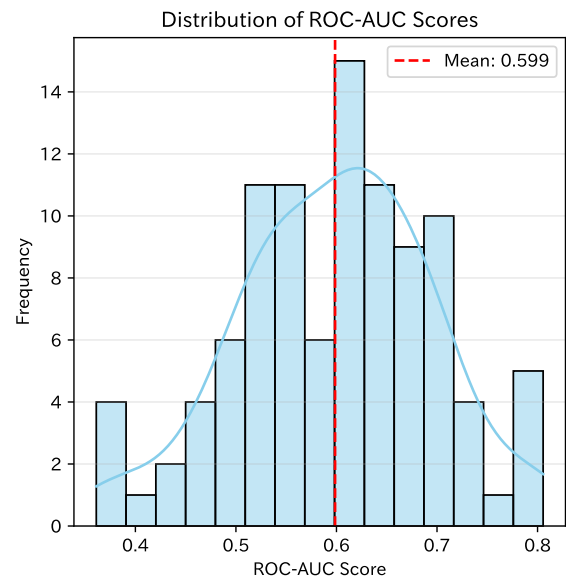


図 2 100 回の試行における ROC AUC スコアの分布

判定においては、同一製品における OPEN 動作の Z スコアと CLOSE 動作の Z スコアのうち、より値が大きい方を採用し、これをその製品の最終的な異常スコアとした。

4.3 検出性能評価

本研究では、異常検知性能の定量的な評価指標として AUC スコアを採用した。評価データの選定による結果のばらつきを定量的に評価するため、100 回の独立した試行を実施した。具体的には、各試行において、ランダムに分割された学習データと評価データを用いてモデルの学習および推論を行い、その都度 AUC スコアを算出した。最終的な評価は、これら 100 個の AUC スコアの平均値および標準偏差を用いて行った。

100 回の試行によって得られた AUC スコアの分布を図 2 に示す。実験の結果、100 試行における平均 ROC AUC スコアは 0.599 となった。この値はランダムな予測を上回ってはいるものの、実用的な異常検知には不十分な水準であり、検出性能は低く留まる結果となった。図 2 を確認すると、試行によって AUC スコアにのばらつきが存在することがわかる。これは、学習に使用される正常データの組み合わせによって、モデルが獲得する正常範囲の境界が変動し、その結果、境界付近に位置する検出難易度の高い異常データの検出可否が変化するためであると考察される。

4.4 異常サンプルごとの検出強度

4.3 において検出性能の低く留まった要因を明らかにするため、異常データごとの Z スコア分布について詳細な分析を行った。図 3 は、異常データ全 6 製品における 100 回試行分の Z スコア分布を箱ひげ図として示したものである。図中の赤色の菱形 (◇) は各製品における 100 試行の

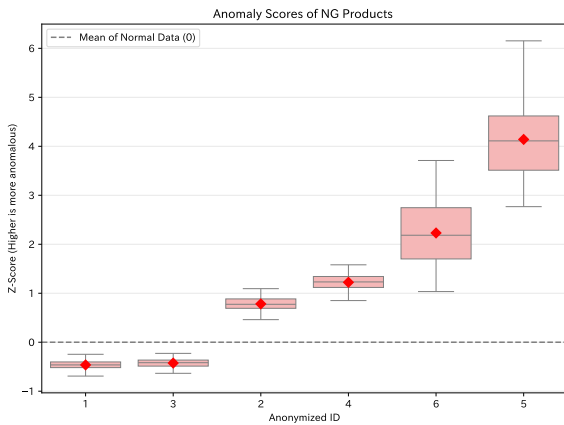


図 3 異常データ全 6 製品における 100 回試行分の Z スコア分布

平均スコアを示している。また、箱の上下端は四分位範囲（第 1 四分位数から第 3 四分位数）を、箱内部の線は中央値を表し、ひげの上下端は外れ値を除くデータの最大値および最小値を示している。なお、横軸は各製品 ID を表しており、図の左側ほど平均 Z スコアが低く、正常データとの判別が困難なサンプルとなるように昇順で並べ替えている。図 3 より、異常データはその Z スコア傾向に基づいて 2 つのグループに大別されることが分かる。図中右側に位置する 4 製品は「高スコア群」であり、閾値を適切に設定することで高い精度で検出できる。それに対し、図中左側に位置する 2 製品は「低スコア群」に該当し、正常データの分布範囲内 ($Z \approx 0$) に重複している、あるいは正常データよりも低い Z スコアとなっている。

検出性能が低くなった原因は、この低スコア群の存在にある。低スコア群の 2 製品については、OPEN/CLOSE いずれのモデルにおいても正常データと類似した再構成誤差しか生じず、結果として標準化された Z スコアも低く算出された。これは、低スコア群の 2 製品の異常が、本手法のオートエンコーダが学習した正常な音響特徴と非常に類似しているか、あるいは背景雑音等の影響により特徴が埋没してしまった可能性が考えられる。

一方で、高スコア群の 4 製品については平均して正常データより高い異常度を出力できており、本手法が特定の種類の異常に対しては有効であることが確認された。今後の課題として、低スコア群のような異常特徴も捉えるための特徴量設計や、より表現力の高いモデル構造への改良が挙げられる。

5. おわりに

本稿では、機械製品の動作音検査において、周期的な異常音を対象とし、エネルギー変動スペクトルと単純化オートエンコーダを用いた教師なし異常検知手法を提案した。本手法では、OPEN 動作と CLOSE 動作の音響特性の違いに着目してそれぞれ個別のモデルを構築するとともに、異常度として Z スコアを採用することで、動作状態に依存し

ない統一的な評価指標を確立した。

実データを用いた 100 回の試行による評価実験の結果、全体としての AUC スコアは 0.599 となった。この結果について詳細な分析を行ったところ、異常データはその検出難易度により「高スコア群」と「低スコア群」の 2 つに傾向が分かれることが確認された。高スコア群については正常データ群から乖離した Z スコアを示しており、提案手法が想定した特徴を持つ異常を安定して検出できることが確認された。一方で、低スコア群については正常データの変動範囲内に Z スコアが留まり、その結果として低い AUC スコアとなった。低スコア群の異常特徴が本手法で抽出した特徴量では捉えられず、正常データの音声との区別が困難であったことを示唆している。

以上の結果から、提案手法は、想定した音響特徴を伴う異常に対しては一定の有効性を有していることが確認された。今後の自動検査システムの導入に向けては、今回捉えられなかった異常特徴も捉えうる高精度な特徴抽出手法の開発が求められる。

参考文献

- [1] Tang, S., Yuan, S. and Zhu, Y.: Cyclostationary Analysis towards Fault Diagnosis of Rotating Machinery, *Processes*, Vol. 8, No. 10, pp.1217:1–15 (2020).
- [2] Wodecki, J., Michalak, A., Hebda-Sobkowicz, J., Wylomańska, A., Szabat, K., Wolkiewicz, M., Pawlak, M. and Zimroz, R.: Cyclostationary Analysis of Vibration Signals from Electric Motor – Understanding of Bi-frequency Map, *Mechanical Sciences*, Vol. 16, pp. 597–614 (2025).
- [3] 井本桂右, 川口洋平: 環境音分析・異常音検知の研究動向, 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 15, No. 4, pp. 268–280 (2022).
- [4] 大瀧 淳, 小菅 太, 三澤 勉, 植原孝之, 藤江龍一: 異音を定量的に検出する手法の開発, 富士通テン技報, Vol. 23, No. 2, pp. 37–46 (2005).
- [5] 笠原太郎, 宗戸俊樹, 坂本湧暉, 高部見好, 伊藤秀浩: 製品動作音の新検査法の開発, *AISIN TECHNICAL REVIEW*, Vol. 28, pp. 39–42 (2024).
- [6] Duman, T. B., Bayram, B. and İnce, G.: Acoustic Anomaly Detection Using Convolutional Autoencoders in Industrial Processes, *Advances in Intelligent Systems and Computing*, Springer, pp. 432–442 (2020).
- [7] Koizumi, Y., Saito, S., Uematsu, H., Kawachi, Y. and Harada, N.: Unsupervised Detection of Anomalous Sound based on Deep Learning and the Neyman-Pearson Lemma, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 27, No. 1, pp. 212–224 (2019).
- [8] Marchi, E., Vesperini, F., Eyben, F., Squartini, S. and Schuller, B.: A Novel Approach for Automatic Acoustic Novelty Detection Using a Denoising Autoencoder with Bidirectional LSTM Neural Networks, *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1996–2000 (2015).
- [9] Primus, P., Haunschmid, V., Praher, P. and Widmer, G.: Anomalous Sound Detection as a Simple Binary Classification Problem with Careful Selection of Proxy Outlier Examples, *Proceedings of the Detection and Classi-*

- fication of Acoustic Scenes and Events 2020 Workshop (DCASE2020)*, pp. 170–174 (2020).
- [10] Kuroyanagi, I., Hayashi, T., Takeda, K. and Toda, T.: Anomalous Sound Detection Using a Binary Classification Model and Class Centroids, *Proceedings of European Signal Processing Conference (EUSIPCO)*, pp. 1995–1999 (2021).
- [11] 松本啓汰, 山田武志, 牧野昭二: 多クラス分類器とその中間層出力の再構成器を用いた異常音検知の検討, *信学技報*, Vol. 122, No. 81, pp. 77–81 (2022).
- [12] Haunschmid, V. and Praher, P.: Anomalous sound detection with masked autoregressive flows and machine type dependent postprocessing, Technical report, DCASE2020 Challenge Technical Report (2020).