

環境騒音調査の音データを対象としたオートエンコーダによる異常音判別

大谷 紀子 研究室

1972007 荒川 永人

1. 背景と目的

環境基本法第 16 条では、人の健康の保護と生活環境の保全のうえで、大気、水、土壌、騒音をどの程度に保つことを目標に施策を実施していくのかを示した環境基準が規定されている。環境騒音調査は騒音に係る環境基準の達成状況を評価するものである。環境騒音調査において対象とする騒音の範囲は道路交通等に起因する騒音であり、等価騒音レベルで評価することが原則となっている。等価騒音レベルは瞬間的な高レベル音に影響されることから、評価の妨げとなる異常音は除外する必要がある。環境省は評価マニュアル[1]において異常音の種類を定義している。除外対象とする異常音は、航空機騒音や鉄道騒音、建設作業騒音、平常ではない自然音、通常は発生しない人工音、測定による付加的な音である。

建設環境研究所は、環境騒音調査において騒音計と IC レコーダを使いデータを記録している。異常音を除外する際には、まず、騒音計から取得した騒音の時刻歴データから異常音の疑いがある箇所を見つける。次に、IC レコーダで録音した異常音の疑いがある箇所の音を聞き、異常音か否かを判断する。以上のような異常音除外作業は多大な労力を要する。本研究では異常音除外作業における人の負担を軽減することを目的として、オートエンコーダに基づく異常音判別手法を提案する。

2. 提案手法

提案手法では環境騒音調査の音データを対象とし、オートエンコーダを用いて異常音を判別する。オートエンコーダはニューラルネットワーク構造

を用いた手法の 1 つであり、入力と出力の差が小さくなるようにニューロンの重みを調整する。したがって、モデルの入力に正常データを用いることで、正常データに含まれる特徴を学習し、学習に用いていないデータに含まれる異常を検知することができる。オートエンコーダは画像データに対する異常検知に利用でき[2]、音データをメルスペクトログラムに変換することで、音データを画像データと同様に扱うことができる。

本手法は、学習フェーズと推論フェーズの 2 段階からなる。データ拡張の際にはノイズ付与、ピッチシフト、時間シフト、音量シフトを行なう。異常度は入力画像と出力画像の平均二乗誤差とする。学習フェーズの手順を以下に記す。

1. 学習データを A と B に分割する。
2. 学習データ A から正常音を取り出す。
3. 取り出した正常音を 1 秒ごとに切り出しデータ拡張によりデータ数を 2 倍にする。
4. 3 で作成した 1 秒ごとの正常音をメルスペクトログラムに変換する。
5. メルスペクトログラムをオートエンコーダに入力し、判別モデル A を作成する。
6. 2 から 5 を学習データ B に対して行い、判別モデル B を作成する。
7. 判別モデル A を使い、学習データ B に異常ラベルを追加する。
8. 判別モデル B を使い、学習データ A に異常ラベルを追加する。
9. 2 から 8 を 2 回繰り返す。
10. 9 により得られた正解ラベルに基づき、学習

データ全体から2から4と同様の手順で正常音のメルスペクトログラムを作成する。

- メルスペクトログラムをオートエンコーダに入力し、判別モデルCを作成する。

異常音に対して網羅的に異常ラベルを付与する作業は非常に困難であることから、用意された正解ラベルは不完全である。したがって、異常音に正常ラベルが付与されている場合は、ラベルを修正する必要がある。7, 8では、対応する学習データに対して異常度を算出し、異常度が閾値を超えている場合に異常ラベルを追加する。閾値は、検証データから抜き出した正常音に対して各モデルが出力する異常度の最大値とする。

次に推論フェーズの手順を以下に記す。

- テストデータを1秒ごとに切り出し、メルスペクトログラムに変換する。
- メルスペクトログラムを判別モデルCに入力し、異常度を求める。
- 異常度が閾値を超えているデータを異常音候補とし、判別モデルCに入力した際の出力画像と入力画像の差分画像を作成する。
- 差分画像のピクセル値の最大値を3.8倍した値から出力画像のピクセル値の平均値を引いた値が正の場合、入力した異常音候補のデータを異常音とする。

オートエンコーダが出力する画像は不鮮明であるため、正常音の異常度が高くなる場合がある。

4により音圧レベルが大きい異常音を異常と判断でき、出力画像の不鮮明さに起因する誤判別を減らすことができる。

3. 評価実験

建設環境研究所が用意した10日分のデータのうち8日分を学習データ、1日分を検証データ、1日分をテストデータとし、ホールアウト検証を

表1:異常音判別結果

適合率	再現率
81%	92%

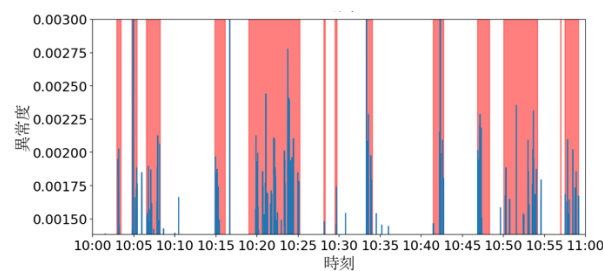


図1:異常度

行なった。学習データに対して学習フェーズ、テストデータに対して推論フェーズの処理を適用し、出力された異常度を可視化した。結果のグラフの一部を図1に示す。赤色の時間帯に異常ラベルが付与されており、異常ラベルが付与された時間帯で異常度が高くなっていることがわかる。正解ラベルに基づき求めた適合率・再現率を表1に示す。建設環境研究所が用意した正解ラベルには、異常音と正常音が交互に混在する時間に対してまとめて異常ラベルが付与されている。したがって、一般的な再現率で評価することができない。本評価実験では異常ラベルが付与された区間のうち、異常と判断されたデータを含む区間の割合を再現率とした。

4. 考察

長時間の音データに対して、異常音の開始・終了時刻を正確に把握し網羅的に異常ラベルを付与することは非常に困難であることから、適合率・再現率から提案手法の真の精度を測ることは難しい。しかし、図1に示すように、概ね異常ラベルが付与されている時間帯で異常度が高くなっており適合率も高いことから、異常度のグラフを見て異常音の疑いがある箇所を特定することができる。したがって、提案手法により、異常音除外作業における人間の負担を軽減できるといえる。

参考文献

- [1] 環境省, “騒音に係る環境基準の評価マニュアル 道路に面する地域編”, 2015.
- [2] H.Nambo, “Anomaly detection with autoencoder”, Journal of Global Tourism Research, Vol.6, No.1, pp.11-14, 2021.