

One-class ニューラルネットワークを用いた腸蠕動音検出 Bowel Sound Detection Based on One-class Neural Network

藤本直希[†] 大野将樹[‡] 岡久稔也[§] 獅々堀正幹[‡]
Naoki Fujimoto Masaki Oono Toshiya Okahisa Masami Shishibori

1. はじめに

近年、腸蠕動音の病態評価を行う方法として、機械学習を用いて、腸蠕動音を検出し評価する研究が行われている。しかし、教師あり学習に用いる訓練データのアノテーションが困難である。なぜなら、腸蠕動音は腸の様々な箇所が蠕動運動することにより発生する音であるため、異なる音の持続時間、発生周期であり、音量も微小であるため、音の数や質を評価することが困難である。そこで、教師なし学習を用いて腸蠕動音の検出を行う。関連研究として One-class SVM を用いた腸蠕動音検出が提案されている。[1] しかし、One-class SVM は前処理部で 5 つの特徴量を抽出し、学習を行うが、最適な特徴量が特定できていない。本研究の目的はオートエンコーダを用いて特徴を自動抽出した後、One-class ニューラルネットワークを用いて腸蠕動音検出の精度を向上させることである。

2. 提案手法

本研究では、オートエンコーダを用いて抽出した特徴量を One-class ニューラルネットワークに入力して分類を行うことで精度の向上を図る。特徴抽出部に用いるオートエンコーダは、腸蠕動音と雑音の両方を学習させ、音声表現の学習を行ったモデルを用いる。その後、識別部に用いるニューラルネットワークには腸蠕動音のみを学習させ、One-class 分類モデルを用いる。

2.1 システムの概要

システムの概要を図 1 に示す。前処理部では、One-class SVM を用いた腸蠕動音検出と同様に、録音したデータからフィルタ処理や雑音除去、有音区間検出などの前処理を行う。特徴抽出部では、オートエンコーダで、特徴抽出を行う。識別部では学習済みのエンコーダ層とニューラルネットワークを用いて One-class ニューラルネットワークで分類を行う。

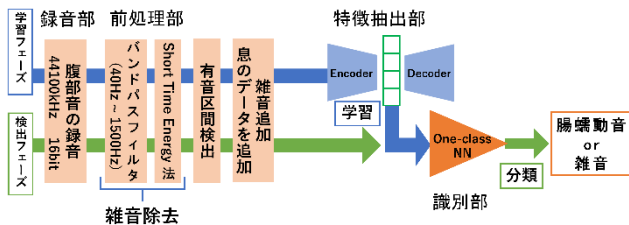


図 1 提案システムの概要

[†] 徳島大学大学院創成科学研究科 Graduate School of Sciences and Technology for Innovation, Tokushima University
[‡] 徳島大学大学院社会産業理工学研究部, Graduate School of Technology, Industrial and Social Science, Tokushima University
[§] 徳島大学大学院医歯薬学研究部, Graduate School of Biomedical Sciences, Tokushima University

2.2 録音部

録音には Cardionics 社の電子聴診器 E-ScopeII を使用し、被験者の臍右側 9cm の位置に当て、Roland 社の R-44 録音機で録音した。録音時のサンプリング周波数は 44100Hz で設定した。西岡ら[2]は多数部位の腸蠕動音を同時測定することにより、腸蠕動音の鳴動回数が最も多かった位置が臍右側 9cm であった。そのため、本研究での測定位置も臍右側 9cm とする。

2.3 前処理部

腹部の録音を行った際、ホワイトノイズが含まれる可能性がある。そこで、腸蠕動音のみを取り出すため、バンドパスフィルタを用いて、腸蠕動音以外の雑音を取り除く。本実験では 40Hz 以上 1500Hz 未満の周波数帯を通過させる設定とした。また、音声信号の強さ STE(Short Time Energy)を用いて、大きすぎる信号を取り除く。その後、音が出ている区間を 0.1 秒ごとに切り出し、切り出したものを 1 サンプルとして定める。

2.4 オートエンコーダ

本研究では、オートエンコーダを用いて、音声信号の特徴抽出を行う。通常のオートエンコーダは全結合層を用いたネットワーク構造であるが、実験では、畳み込み層やプーリング層を用いた CAE(Convolutional AutoEncoder)と時系列データを考慮して GRU-AE(Gated Recurrent Unit AutoEncoder)[3]を用いた特徴抽出を行う。CAE のエンコーダ層には 1 次元の畳み込み層と 1 次元のプーリング層を 3 層ずつ、デコーダ層には 1 次元の畳み込み層と 1 次元のアップサンプリング層を 3 層ずつ組み合わせたネットワーク構造で訓練データを用いて、バッチサイズ 1、エポック数 500 で学習する。GRU-AE は全結合層を GRU に置き換えたネットワーク構造を持つオートエンコーダである。1 層のエンコーダ層と 1 層のデコーダ層を組み合わせたネットワーク構造で訓練データを用いて、バッチサイズ 128、エポック数 300 で学習する。

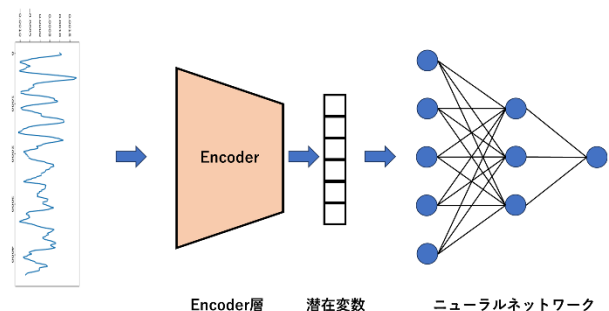


図 2 One-class ニューラルネットワーク

2.5 One-class ニューラルネットワーク

本研究では、One-class ニューラルネットワーク[4]を用いて、腸蠕動音と雑音の分類を行う。One-class ニューラルネットワークとは、オートエンコーダとニューラルネットワークを組み合わせて、正常なデータ(正例)のみの特徴を学習し、異常なデータ(負例)を検出する異常検知手法である。図 2 に One-class ニューラルネットワークの概要を示す。オートエンコーダの潜在変数が入力データの特徴を抽出できる点を利用して、正常なデータのみを使用する教師なし学習である。実験では腸蠕動音を正例として One-class ニューラルネットワークを学習し、腸蠕動音と雑音の 2 値分類を行う。学習に用いた損失関数を式 (1) に示す。

$$L = \min_{w,V,r} \frac{1}{2} \|w\|_2^2 + \frac{1}{2} \|V\|_F^2 + \frac{1}{v} \cdot \frac{1}{N} \sum_{n=1}^N \max(0, r - \langle w, g(VX_n) \rangle) - r \quad (1)$$

w と V はニューラルネットワークの重み、 r は原点との距離、 $\| \cdot \|_2$ はユークリッドノルム、 $\| \cdot \|_F$ はフロベニウスノルム、 X_n は入力データ、 $g(\cdot)$ は中間層の活性化関数、 N は全訓練データ数で $v \in (0, 1)$ はソフトマージンの許容度を表している。 r の更新式を式 (2) に示す。 \hat{y}_n は出力を表している。

$$\arg \min_r \left(\frac{1}{Nv} \cdot \sum_{n=1}^N \max(0, r - \hat{y}_n) \right) - r \quad (2)$$

3. 実験

本実験の目的は、提案手法の有効性を示すことである。

3.1 実験環境

本実験では、静音環境下において 1 名の被験者から電子聴診器を用いて就寝中に腹部の音を約 6 時間録音した。約 6 時間の音声から、先頭 1 時間と末尾 1 時間は雑音が多いためあらかじめ除き、約 4 時間分の音声に対して、前処理を行う。前処理を通過し、残ったサンプルを訓練データと推論データとして 8:2 に分ける。また、推論データは腸蠕動音のみのデータに、雑音のデータを加えて分類を行う。雑音には、ESC-50 データセット[5]から Breathing カテゴリのデータを使用し、腸蠕動音の 1 サンプルごとに音声信号の最大値を取得し、全サンプルの最大値から中間値をとり、その中間値と同程度の音声信号になるように雑音の信号に調整を加えた。

3.2 評価指標

分類精度の評価指標に、感度と特異度を用いて評価を行う。感度が高いほど、腸蠕動音を見逃すことなく検出することができ、特異度が高いほど、雑音の誤検知が少ないと言える。一般的に医学分野では、陽性を陰性と誤識別することが問題とされるため感度が高いことが重要である。

3.3 実験結果

One-class SVM との精度比較を表 1 に示す。提案手法では、One-class SVM の精度より向上する結果は見られなかった。CAE を用いた One-class ニューラルネットワークは、One-class SVM と同様な精度が見られ、GRU-AE を用いた One-class ニューラルネットワークは、感度が低く、特異度が高い結果となった。

表 1 実験結果

	従来手法	CAE	GRU-AE
感度	97.3%	96.1%	38.9%
特異度	75.1%	64.7%	80.5%

3.4 考察

提案手法である CAE と GRU-AE を用いて One-class ニューラルネットワークで腸蠕動音の分類を行うと、感度は CAE を用いた One-class ニューラルネットワークでの分類精度の方が高いことが確認できる。しかし、再構成誤差は表 2 からわかるように GRU-AE の方が低いため、GRU-AE の方が特徴量を抽出できているのではないかと考える。

表 2 オートエンコーダによる再構成誤差

	CAE	GRU-AE
損失関数	7.25×10^{-7}	7.70×10^{-9}

また、分類精度が低い原因としてデータに問題があるのではないかと考えられる。本研究に用いたデータは就寝時に取られており、寝返りで服がすれるなど、除かれていない雑音があるため、雑音も腸蠕動音の特徴としてオートエンコーダで学習され、雑音を腸蠕動音と誤分類するため、分類精度の低下につながったのではないかと考える。録音を静音環境で行うことで高品質なデータを作成し、訓練データを改善することで分類精度が向上すると期待できる。

4. おわりに

本研究では、自動で特徴抽出を行うオートエンコーダを用いて One-class ニューラルネットワークで腸蠕動音と雑音の分類を行う手法を提案した。同じ教師なし学習手法である従来手法の One-class SVM を比較対象として実験を行ったが、実験結果より分類精度の向上は見られなかった。CAE を用いた One-class ニューラルネットワークでは、従来手法と同様な分類精度が見られたが、GRU-AE を用いた One-class ニューラルネットワークでは、感度が低く、特異度が高いという結果であった。今後の課題として、録音環境を見直し、訓練データに雑音が入っていない腸蠕動音データなど、高品質なデータセットを用意することで、精度の向上が期待される。

謝辞

本研究は JSPS 科研費 JP23K11896 の助成を受けたものです。

参考文献

- [1] 大野将樹, 森内翔希, 獅々堀正幹, 岡久稔也 “One-class SVM に基づく教師なし腸蠕動音検出”, 電気・電子・情報関係学会四国支部連合大会講演論文集, Vol.2022, pages.ROMBUNNO.14-1 (2022).
- [2] 西岡潤司 他, “Multi point bowel-sound analysis system(MPBAS)の試作による腸蠕動音測定部位の最適化”. 人工臓器(人工臓器学会), Vol.47, pages. S.166 (2018).
- [3] Kyunghyun Cho et al. “Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”, arXiv, 1406.1078 (2014).
- [4] Raghavendra Chalopathy et al. “Anomaly Detection using One-Class Neural Networks”, arXiv, 1802.06360 (2019).
- [5] K. J. Piczak. “ESC: Dataset for Environmental sound Classification”, Proceedings of the 23rd Annual ACM Conference on Multimedia, pages. 1015-1018 (2015).