

令和 2 年 6 月 8 日現在

機関番号：10103

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00373

研究課題名(和文)連鎖パタンマイニングを用いた心電図異常検出

研究課題名(英文)Automatic detection of abnormal ECG based on linkage pattern mining

研究代表者

岡田 吉史 (Okada, Yoshifumi)

室蘭工業大学・大学院工学研究科・准教授

研究者番号：00443177

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：本研究の目的は、心電図データにおける異常な波形領域を自動検出する連鎖パタンマイニング手法を開発することであった。まず、既存の連鎖パタンマイニング手法を改良し、より高速かつ高精度な手法を開発した。実験として、改良法の実装プログラムを実際の心電図データに適用した結果、正常/異常な波形領域を適切に区別して視覚的に提示できることがわかった。さらに、本手法により得られた正常/異常波形は、機械学習を用いた識別モデルの構築において、有効な訓練セットとして利用可能であることが示された。

研究成果の学術的意義や社会的意義

本研究で開発した技術は、医療現場のスタッフが心電図を用いて心疾患を迅速に診断するための有効な支援ツールとなりえる。また、既存の機械学習を用いた心電図解析では正常/異常の訓練データは手作業で収集されており、この作業には専門的な知識と多大な時間が必要とされていた。一方、本研究で開発した技術は、正常/異常な心電図波形を自動で高速に特定できるため、今後の心電図解析研究を大きく加速・進展できると期待される。

研究成果の概要(英文)：The aim of this study was to develop a linkage pattern mining method for automatically detecting abnormal waveform region in ECG data. First, the existing linkage pattern mining method was improved to a faster and more accurate method. In an experiment of applying a program implementing the improved method to real ECG data, it was shown that normal/abnormal ECG regions were partitioned adequately and displayed visually. Furthermore, it was presented that the normal/abnormal ECG waveforms can be available as an effective training dataset in classification model construction based on machine learning.

研究分野：データマイニング、機械学習

キーワード：心電図 連鎖パタンマイニング 畳み込みニューラルネットワーク

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年、機械学習アルゴリズムを用いて、心電図データから心疾患の識別を行う研究が数多くなされている。それらは、事前に心電図の異常波形を手動または半手動で収集し、識別モデルを学習するアプローチをとっている。しかし、膨大な心電図データから異常波形のみを手作業で収集する作業は、専門的な知識が必要とされるだけでなく多大な時間を要する。現在、PhysioBankのような公共データベースにおいて、様々な種類の心疾患の心電図データが日々蓄積されつつある。今後も、それらのデータを利用して、種々の心疾患を識別する手法が提案されるものと考えられる。ゆえに、心電図から異常波形を自動的に検出する技術は、今後の心電図自動解析研究を大きく加速させ、さらなる進展へと導く上で極めて重要になってくると考える。

2. 研究の目的

本研究課題の目的は、異なる部位(以下、誘導)から得られる複数の心電図データから、異常波形を自動検出する手法を開発することである。異常波形の検出手法として、以前申請者らが開発した連鎖パターンマイニング手法を採用する。また、連鎖パターンマイニングにより検出された異常波形の妥当性を調べるため、実際の心疾患の心電図データから検出された異常波形を用いて識別モデルを構築し、識別精度の評価を試みる。

3. 研究の方法

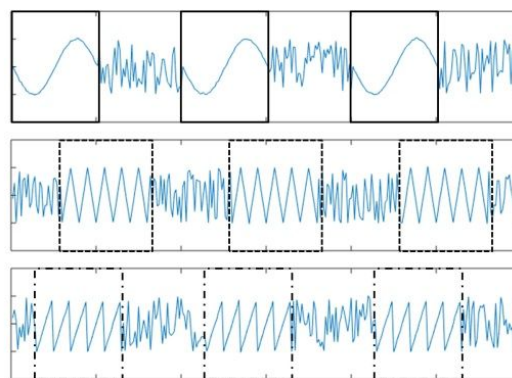
本研究課題は、(1)連鎖パターンマイニングによる心電図異常検出、および(2)畳み込みニューラルネットワークによる心疾患識別の2つのステップで構成される。以下、それぞれの概要について説明する。

(1) 連鎖パターンマイニングによる心電図異常検出

連鎖パターンマイニング手法は、図1に示すように複数の時系列データに跨って繰り返し出現するパタン群(連鎖パタン)を発見するものである。正常な心電図波形は、複数誘導に跨って規則的かつ周期的に出現する。この性質を利用して、まず連鎖パターンマイニングによって複数時系列に渡って反復する正常波形群(すなわち、連鎖パタン)を特定する。続いて、それらの正常波形群をマスキングすることで、異常波形領域を検出する。

(2) 畳み込みニューラルネットワークによる心疾患識別

心疾患の識別モデルを構築するため、訓練データとして、連鎖パターンマイニングにより検出された異常波形領域の画像群(異常クラス)と、正常な心電図から切り出された正常波形の画像群(正常クラス)を準備する。これらを畳み込みニューラルネットワークで学習し、識別モデルを構築する。連鎖パターンマイニングで得られた異常波形の妥当性を調べるため、実際の心疾患(心室期外収縮)の心電図データを用いた識別テストを実施する。






連鎖パタン : {  ,  ,  }

図1 連鎖パターンマイニング

4. 研究成果

(1) 連鎖パターンマイニングの高速化

連鎖パターンマイニングは、各時系列データから繰り返し出現するパタン(以下、頻出パタン)を抽出する、異なる時系列データを跨ぐ頻出パタン間の連鎖をチェックする、という2つのステップで行われる。既存の連鎖パターンマイニング手法(以下、既存法)は、1つ目のステップにおいて、離散化されたデータを対象として離散データ列の組合せ的なチェックを行っていたため、膨大な計算時間を要していた。そこで本研究では、時系列データの開始点からシフトして得られる部分系列をキーとして頻出パタンを探索するアルゴリズムを新たに開発した。この方法は、離散データ列に関する組合せ探索を行う必要がないため、高速な頻出パタン抽出が可能とな

る。実験として、2種類の人工データを用いて計算時間を評価した結果、いずれのデータセットにおいても10%程度にまで計算時間を削減できることがわかった。

(2) 連鎖パターンマイニングの高精度化

上述のように連鎖パターンマイニングは、2つのステップで実行される。既存法では、2つ目のステップにおいても、2つの重大な問題点があった。1つ目は、異なる時系列におけるパタン間の出現順序を考慮していなかった点である。これにより、連鎖パターンを構成するパタンの種類は同じだが、それらの出現順序が異なるパタン群を同一の連鎖パターンとしてみなしていた。2つ目は、異なる時系列のパタン間で時間的な重複区間がなければ連鎖パターンとはみなさなかった点である。これにより、異なる時系列におけるパタンが互いに連鎖的な出現傾向を示していても、時間的な重複区間が無い場合は連鎖パターンとみなされず無視されていたことになる。

そこで本研究では、上記の問題を解決するためにアルゴリズムの改良を行った。実験として、ランダムデータに人工的に連鎖パターンを埋め込んだ7つの人工データセット(Dataset1 ~ Dataset7)を用いて、連鎖パタンの抽出精度を評価した。図2は、7つのデータセットに関する埋め込み連鎖パタンの抽出精度(PrecisionとRecall)を示している。この図において、既存法のDataset1とDataset2で抽出精度が示されていないのは、埋め込み連鎖パターンが全く抽出されなかったことが理由である。図からわかるとおり、全てのデータセットにおいて、改良法の抽出精度は、既存法のそれを大きく上回っていることがわかる。

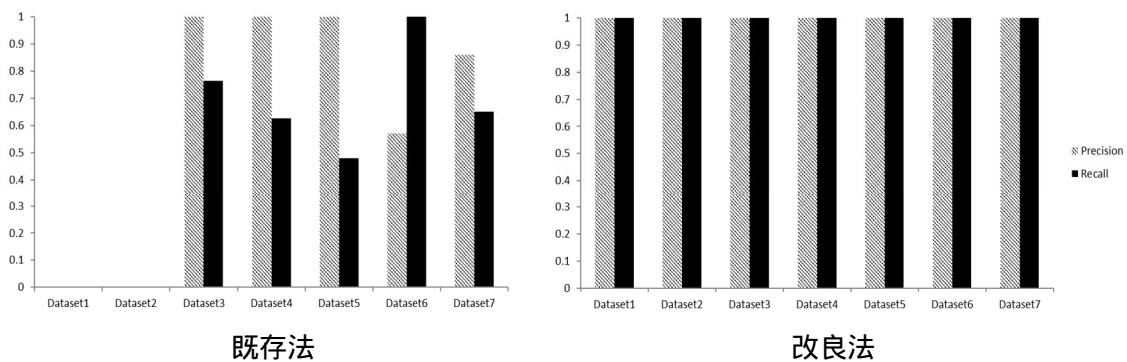


図2 既存法と改良法における連鎖パタンの抽出精度

(3) 連鎖パターンマイニングを用いた心電図異常検出

既存法を高速化・高精度化した改良法を用いて、心電図データにおける異常波形を検出するプログラムを開発した。このプログラムは、心電図データの前処理、連鎖パターンマイニング、連鎖パターンマスキング、結果出力の4つのサブプログラムで構成される。図3は、健康者の正常な心電図(Healthy ECG)と心筋梗塞患者の心電図(Disease ECG)を入力として、本プログラムを実行した結果である。この図において、黄色の領域が正常と判定された波形領域(すなわち、連鎖パターンとして特定された領域)、青色の領域が異常と判定された波形領域(すなわち連鎖パターン以外の領域)である。この図からわかるように、正常な心電図では、全ての領域が本プログラムによって正常と判定されていることがわかる。逆に、心筋梗塞の心電図では、所々異常と判定されている部分があることがわかる。

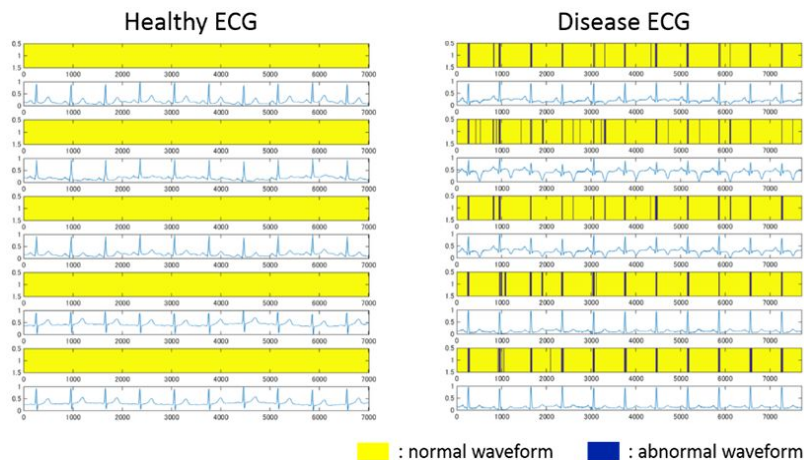


図3 心電図異常検出プログラムの実行例

(4) 連鎖パターンマイニングと畳み込みニューラルネットワークに基づく心疾患識別実験

この実験の目的は、改良された連鎖パターンマイニング手法によって検出した異常波形の妥当性を評価することである。本研究では、心室期外収縮と呼ばれる不整脈患者の心電図を対象として実験を行った。心室期外収縮とは、心臓の心室異常により、正常波形の反復の中に余分な異常波形が出現する不整脈である。まず、心室期外収縮を持つ複数患者の心電図データに連鎖パターンマイニングを適用し、異常波形領域を検出した。続いて、検出した異常波形領域から一定区間の画像を切り出して保存し、これを異常クラスと定義した。また、正常な心電図から、異常クラスの画像と同サイズの正常波形画像を切り出して保存し、これを正常クラスと定義した。続いて、2次元畳み込みニューラルネットワークを用いて、これら2クラスの画像分類を行うための識別モデルを構築した。本実験では、3分割交差検定により、心室期外収縮か否かの識別が適切に行われるかどうかにより、連鎖パターンマイニングによる心電図異常検出の妥当性を評価した。識別精度指標として、Accuracy, Sensitivity, Specificity を用いた。表1は、本実験 (This study) および先行研究 (Mohamed et al., 2016, Yonghan et al., 2017) で得られた心室期外収縮の識別精度である。この表から、全ての指標において、先行研究よりも高い精度を示していることがわかる。この結果は、心室期外収縮を特徴づける異常波形が、連鎖パターンマイニングにより適切に検出されたことを示唆している。

表1 心室期外収縮の識別精度

Author (Year)	Techniques/Features	Results
Mohamed et al., 2016	· Fractional linear prediction · k-NN	Accuracy = 95.0%
		Sensitivity = 95.0%
Yonghan et al., 2017	· Discrete wavelet transforms	Specificity = 85.0%
		Accuracy = 99.0%
		Sensitivity = 94.4%
This study	· Linkage pattern mining · CNN	Specificity = 99.8%
		Accuracy = 99.9%
		Sensitivity = 99.7%
		Specificity = 99.9%

< 引用文献 >

Mohamed Lamine Talbi and Philippe Ravier, "Detection of PVC in ECG signals using fractional linear prediction", Biomedical Signal Processing and Control, Vol. 23, pp.42-51, 2016.

Yonghan Jung and Heeyoung Kim, "Detection of PVC by using a wavelet-based statistical ECG monitoring procedure", Biomedical Signal Processing and Control, Vol.36, pp.176-182, 2017.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Kaiji Sugimoto, Yudai Kon, Saerom Lee, and Yoshifumi Okada	4. 巻 178
2. 論文標題 Detection and localization of myocardial infarction based on a convolutional autoencoder	5. 発行年 2019年
3. 雑誌名 Knowledge-Based Systems	6. 最初と最後の頁 123-131
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.knosys.2019.04.023	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Saerom Lee, Kaiji Sugimoto, and Yoshifumi Okada	4. 巻 46
2. 論文標題 Linkage Pattern Mining using Interval and Order of Pattern Appearance	5. 発行年 2019年
3. 雑誌名 IAENG International Journal of Computer Science	6. 最初と最後の頁 691-698
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計10件（うち招待講演 0件 / うち国際学会 3件）

1. 発表者名 小林哲平, 金曼大, 岡田吉史
2. 発表標題 心電図データを用いた心室期外収縮の自動検出
3. 学会等名 生命ソフトウェア・感性工房 合同シンポジウム2019
4. 発表年 2019年

1. 発表者名 金曼大, 岡田吉史
2. 発表標題 畳み込みニューラルネットワークを用いた心筋梗塞の検出および梗塞部位の特定
3. 学会等名 生命ソフトウェア・感性工房 合同シンポジウム2019
4. 発表年 2019年

1. 発表者名 牧野晃希, 金憂大, 岡田吉史
2. 発表標題 心電図のR-R間隔画像への2次元畳み込みニューラルネットワークの適用と致死性不整脈の識別
3. 学会等名 生命ソフトウェア・感性工房 合同シンポジウム2019
4. 発表年 2019年

1. 発表者名 Kaiji Sugimoto, Saerom Lee and Yoshifumi Okada
2. 発表標題 Discrimination of ECG Abnormality based on a Normal ECG Wave Model Implementing a Denoising Model
3. 学会等名 IMECS2019 (国際学会)
4. 発表年 2019年

1. 発表者名 李セロン, 杉本階嗣, 岡田吉史
2. 発表標題 パタン出現の間隔と順序を考慮した連鎖パタンマイニングに関する研究
3. 学会等名 第14回日本感性工学会春季大会
4. 発表年 2019年

1. 発表者名 米谷柊太郎, 杉本階嗣, 李セロン, 岡田吉史
2. 発表標題 実数値の系列データを対象としたノイズに頑健な頻出パタンマイニング手法
3. 学会等名 生命ソフトウェア・感性工房・而立の会合同シンポジウム2018
4. 発表年 2018年

1. 発表者名 杉本階嗣, 李セロン, 岡田吉史
2. 発表標題 Convolutional Autoencoderを用いた心電図波形モデルの構築と病気診断への応用
3. 学会等名 生命ソフトウェア・感性工房・而立の会合同シンポジウム2018
4. 発表年 2018年

1. 発表者名 Kaiji Sugimoto, Saerom Lee and Yoshifumi Okada
2. 発表標題 Deep Learning-based Detection of Periodic Abnormal Waves in ECG Data
3. 学会等名 IMECS2018 (国際学会)
4. 発表年 2018年

1. 発表者名 岡田吉史, 李セロン
2. 発表標題 連鎖バタマイニングの心電図データへの適用
3. 学会等名 第13回日本感性工学会春季大会
4. 発表年 2018年

1. 発表者名 Saerom Lee, Kaiji Sugimoto and Yoshifumi Okada
2. 発表標題 Detection of Abnormal ECG Waveform Based on Linkage Pattern Mining
3. 学会等名 ICBAKE2017 (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----