

卒業研究論文

論文題目

付け爪センサによる生体信号を用いた深層学習に基づく心拍推定

提出年月日		3	令和 5 年 2 月 22 日	
学科		科	電気情報工学科	
氏		名	島田 優斗	印
指導教員(主査)		査)	北村 大地 講師	印
副	副 查 雛元 洋一 助教 印		ED	
学	科	長	辻 正敏 教授 印	

香川高等専門学校

Heart Rate Estimation Based on Deep Learning Using Biosignal measured by Nail Tip Sensor

Yuto Shimada

Department of Electrical and Computer Engineering National Institute of Technology, Kagawa College

Abstract

Remote monitoring is important in a home health care for handling sudden changes of patient's physical conditions. One approach to achieve remote monitoring is that patients wear biometric sensors at all times. However, since almost all biometric sensors are used in contact with the body surface, and wearing them for a long period causes a decrease in quality of life and hygiene problems. As another approach to solve this problem, a nail tip sensor was developed. This sensor is a wearable device that is attached to the surface of the fingernail. Therefore, there is no stress about wearing it, and it can be firmly fixed for a long time. This sensor can measure photoplethysmography (PPG) by observing the intensity of reflected light associated with the dilation and contraction of the blood vessels. Thus, monitoring the heart rate in daily life can be realized. Nevertheless, the PPG derived from the heart rate obtained from the nail tip sensor is so weak, and it contains enormous noise caused by body movements and other factors. To cope with this problem, fundamental frequency estimation algorithms for periodic signals were utilized to estimate the heart rate from PPG signals. Although these methods achieved accurate heart rate estimation during sleeping, it was difficult to estimate the heart rate during daily activities. In this thesis, I confirm the efficacy of a supervised heart rate estimation method using the nail tip sensor and a deep neural network (DNN) to realize even more accurate heart rate estimation. First, I develop a labeling application to easily generate correct heart rate values for the measured PPG. Using the labeled data prepared by this application, I build several types of DNN models and experimentally investigate their accuracy in predicting heart rate from PPG. As a result, it was revealed that the bidirectional long-short term memory model can estimate the heart rate with high accuracy both during sleeping and during daily activities.

Keywords: nail tip sensor, heart rate estimation, deep neural network, recurrent neural network

在宅医療における遠隔モニタリングは、患者の体調の急な変化に対応するために重要であ る.この遠隔モニタリングを実現する一つの方法として,患者に生体センサを常時装着させる ことが考えられる、しかしながら、多くの生体センサは体表面への接触を伴うため、長時間 や常時の装着は生活の質の低下や衛生面の問題が生じてしまう.この問題を解決する1つの アプローチとして,付け爪型センサが現在開発されている.このセンサは爪の表面に装着し 固定するウェアラブルデバイスであり,不快な装着感がなく長期間にわたる強固な固定が可 能である.このセンサでは,血管の拡張・収縮に伴う反射光の強弱変化から,光電脈波信号 (photoplethysmography: PPG) を得ることができるため,日常生活における心拍のモニタリ ングに応用することが期待されている.ただし,付け爪センサから得られる心拍由来の PPG は微弱であるため,体動等によって大きなノイズが混入する問題がある.これまでに,周期信 号の基本周波数を推定する手法を付け爪センサの PPG に適用する教師無し心拍推定法が研究 された、その結果、睡眠時の心拍の高精度な推定は可能となったが、通常活動時の推定は困 難であった.本論文では,更なる高精度な心拍推定の実現を目指して,付け爪センサと深層 ニューラルネットワーク (deep neural network: DNN) を組み合わせた教師あり心拍推定法 について有効性を確認する.まず,測定した PPG に対して正解の心拍値を簡便に付与するた めのラベリングアプリケーションを開発する.このアプリケーションによって用意されたデー タを用いて、複数種類の DNN モデルを学習し、PPG から心拍を予測する精度について実験 的に調査する.実験結果から,双方向の再帰構造を持つ長・短期記憶型の DNN モデルが,睡 眠時及び通常活動時の両方において高い精度で心拍を予測できることを確認した.

目次

第1章	緒言	1
1.1	本論文の背景	1
1.2	本論文の目的	3
1.3	本論文の構成	4
第2章	付け爪型センサの概要	6
2.1	まえがき	6
2.2	特徴	6
2.3	動作原理	7
2.4	測定波形	7
2.5	本章のまとめ	9
第3章	基礎技術	13
3.1	まえがき	13
3.2	STFT	13
3.3	MLP に基づく DNN	14
3.4	RNN に基づく DNN	15
	3.4.1 LSTM	17
	3.4.2 GRU	18
3.5	BiRNN に基づく DNN	19
3.6	本章のまとめ	20
第4章	提案手法	21
4.1	まえがき	21
4.2	提案手法の動機	21
4.3	ラベリングアプリケーションの開発	23
4.4	PPG の前処理	24
4.5	DNN モデルの入力	28
4.6	DNN モデルの出力と学習	28
47	太音のまとめ	29

第5章	心拍の予測実験	30
5.1	まえがき	30
5.2	評価指標	30
5.3	実験条件	31
5.4	実験結果	32
	$5.4.1$ データ 2 のみを用いた学習と評価 \dots	33
	5.4.2 全データを用いた学習と評価	34
5.5	本章のまとめ	35
第6章	結言	51
謝辞		52
参考文献		52
付録 A	データ1のみを用いた学習と評価	55
付録 B	データ3のみを用いた学習と評価	62
付録 C	全データを用いた学習とデータ2及びデータ3に対する評価	69

第1章

緒言

1.1 本論文の背景

在宅医療において、患者の体調や健康状態を管理することは重要である.例えば遠隔で心拍や血圧等を常時モニタリングすることができれば、急な体調変化にもすぐに対処できる可能性が高まる.また、健常な人に対しても、健康状態を常時モニタリングすることで、心筋梗塞や不整脈等のような突発的な病気に対して素早い処置が可能になると思われる.このような理由から、心拍や血圧等の生体情報を手軽にモニタリングできるウェアラブルデバイス [1, 2] やシステムが重要視されている.現在は、Fig. 1.1 のようなウェアラブルデバイスが販売されており、これらを装着することで遠隔での心拍等のモニタリングが可能となっている.しかしながら、これらのデバイスは入浴及び睡眠時に着脱の必要があるといった煩わしさに加えて、皮膚表面の蒸れに起因する衛生面の問題や接触皮膚炎の誘発等の問題が生じる.

そこで,現行のデバイスの問題の対策として付け爪型センサの開発が行われている [3].このデバイスは Fig. 1.2 のように,爪の表面に装着し固定するウェアラブルデバイスであり,爪の下にある血管の拡張・収縮に起因する微小な爪表面の反射光の変化を計測するものである.デバイス内部の受光センサから得られる信号は光電脈波信号 (photoplethysmography: PPG) と呼ばれ,心拍等の情報を含む生体信号として利用できる.従って,PPG を解析することで,爪の表面に装着したデバイスから心拍を推定することが可能である.

付け爪型センサの利点は以下の3点である.

- 爪には感覚神経が無いため, 不快な装着感が無い
- 爪は組織が固いため,数週間固定を維持できる
- 複数の爪に装着することで多点同時計測ができ、より発展的な計測が期待できる

しかしながら,爪から観測される PPG は体動や外的要因によるノイズが多く混入してしまう問題がある.このため,ノイズに頑健な心拍推定法が必要となる.

各種センサから心拍を推定する手法はこれまでに様々な研究が行われており,これらは Fig. 1.3 のように分類される.心電信号 (electrocardiograph: ECG) から推定する手法として カルマンフィルタを用いる手法 [4],PPG から推定する手法として 2 種類の適応フィルタを





Fig. 1.1. Wearable devices used for remote monitoring [1, 2].

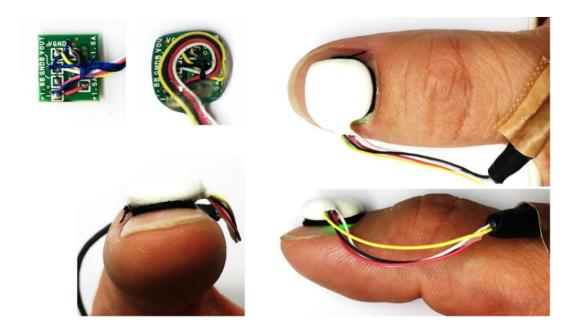


Fig. 1.2. Nail tip sensor for PPG measurement.

用いる手法 [5],及び深層学習を用いる手法 [6,7] などが提案されている.そのほかにも,ビデオから推定する手法 [8,9,10] や音声信号から推定する手法 [11] などが提案されている.また,付け爪型センサにおいては,音楽信号処理における基本周波数推定法を用いた心拍推定精度の検討が行われた [12,13].この手法では,正規化自己相関関数法 [14],ケプストラム法 [15],YIN 法 <math>[16],及び MUSIC 法 [17] の 4 つの教師無し手法が適用された.実験で明らからになったこととして,いずれの手法においても,睡眠時のような安静にしている状態では心拍を推定することができるが,日常生活下の通常活動時になると高精度な心拍推定が困難になる

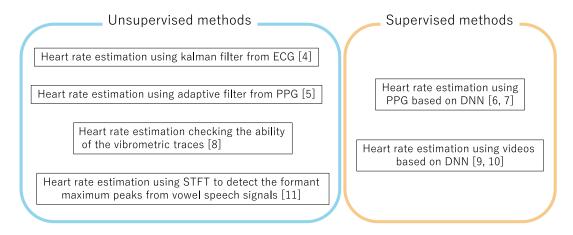


Fig. 1.3. Summary of heart rate estimation methods.

という結論となった.本論文では,付け爪型センサにおいて通常活動時も高精度に心拍を推定 することを目指し,教師ありの手法の適用について検討する.

1.2 本論文の目的

付け爪型センサが提案された文献 [3] では,心拍を推定する手法としてゼロクロス法が用いられた.ゼロクロス法の概要を示した図を Fig. 1.4 に示す.ゼロクロス法とは,正負が反転している時間は計信号の間隔を信号の基本周波数として求める手法である.この手法では,正負の反転を検知しない区間であるデッドタイムを定めるが,これを手動で決定しなければならないという欠点がある.そこで,よりノイズに頑健な心拍推定法を目指して音楽信号処理における基本周波数推定法 [12, 13] を用いる心拍推定法が提案された.音楽信号解析分野において代表的な 4 つの手法である正規化自己相関関数法 [14],ケプストラム法 [15],YIN 法 [16],MUSIC 法 [17] がそれぞれ適用された.特に,MUSIC 法は,観測信号に含まれる白色ノイズに対して頑健な基本周波数推定法であるが,前述の通りいずれの手法においても,通常活動時の高精度な心拍推定は困難であった.

本論文では、付け爪型センサで得られる PPG に対して、体動等のノイズが多く混入する通常活動時においても高精度に心拍を推定することを目指し、教師ありの心拍推定手法の活用を提案する。具体的には、Fig. 1.5 に示すように、PPG を短時間フーリエ変換(short-time Fourier transform: STFT)して得られるスペクトログラムを入力し、心拍値を出力する深層ニューラルネットワーク(deep neural network: DNN)モデルを学習し、その精度について実験的に評価する。この DNN の学習には、入力データとその正解心拍値(ラベル)のペアが必要となるが、現状測定されている PPG には信頼のおけるラベルが付与されていない問題がある。そこで本研究では、まず PPG に対して簡便に心拍のラベルを付与できるアプリケーションを開発する。次に、このアプリケーションを利用して作成された学習データを用いて、多層パーセプトロン(multilayer perceptron: MLP)、長・短期記憶(long short-term memory:

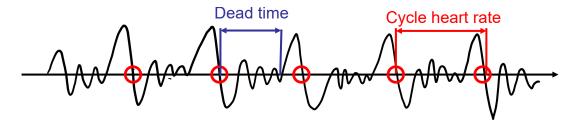


Fig. 1.4. Overview of the zero-cross method.

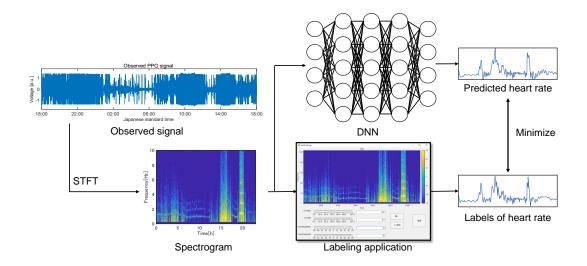


Fig. 1.5. Proposed approach for estimating heart rate based on DNN.

LSTM) ユニット , ゲート付き回帰型ユニット (gated recurrent unit: GRU), 双方向再帰型 LSTM (bidirectional LSTM: BiLSTM) ユニット , 双方向再帰型 GRU (bidirectional GRU: BiGRU) の 5 種類の DNN モデルを構築する. 最終的に , 各 DNN モデルの心拍推定精度に ついて実験的に調査し , 付け爪型センサに対して教師あり手法を適用することの妥当性と , 適切なモデルについて議論する.

1.3 本論文の構成

2章では,本論文で扱う PPG を測定する付け爪型センサの概要について述べる.付け爪型センサの利点と動作原理を示し,実際に付け爪型センサによって測定された波形を掲載する. 3章では本研究において重要となる STFT,MLP,再帰型ニューラルネットワーク(recurrent neural network: RNN),及び双方向 RNN(bidirectional RNN: BiRNN)の 4 つの技術について説明する.4章では,教師ありの心拍推定を行うにあたって必要な正解値を付与するためのアプリケーションについて説明する.その後,DNN の入出力及び損失関数について説明す

る.5章では測定波形と作成した教師データを用いて学習モデルを作成し,心拍推定の精度を 検証する.6章では,本論文の結論を述べる.

第2章

付け爪型センサの概要

2.1 まえがき

本章では、本研究で用いる付け爪型センサ [3] の概要を説明する.まず 2.2 節では、付け爪型センサの特徴について詳細に述べる.ここでは、付け爪型センサの開発の重要性や、本論文で問題とされている外乱の特徴についても説明する.2.3 節では、付け爪型センサの計測原理について述べる.計測原理では、問題視されている外乱の発生原理についても詳しく解説する.2.4 節では、付け爪型センサから得られた測定波形について述べる.この測定波形が本研究で取り扱う入力信号となる.最後に、2.5 節では、本章のまとめを述べる.

2.2 特徴

在宅医療において、患者の心拍をリアルタイムで遠隔モニタリングすることは必須である、しかし、心拍測定を目的として一般的に用いられるセンサは、身体に取り付ける電極によって、不快な装着感や着脱のストレス、及び衛生面等の多くの問題が生じる.一方で、文献 [3] で提案されている付け爪型センサでは、爪には感覚神経が無いことや、組織が固く長期間に渡る固定が維持できることから、一般的に用いられているセンサに生じる問題を全面的に解消できる.なお、本センサでは、センサの固定のためにジェルネイルを用いているため、数週間程度の固定も実現できる.さらに、人間の爪は両手両足で合計 20 枚存在するため、将来的には複数の爪に装着することで多点同時計測も可能となり、より正確にモニタリングすることができると考えられる.これが実現できれば、より発展的な計測も期待でき、心拍測定のみでなく、血圧測定にも応用できると予想される.

血液は心拍に合わせて拡張や収縮を周期的に繰り返しながら血管を流れている.そこで,本センサでは,爪が皮膚に接触している面(爪床)にある毛細血管の拡張及び収縮から生じる微小な色の変化を調べることで心拍を推定している.ところが,爪周辺の皮膚は非常に繊細であるため,体動や筋動に起因する外乱からも,色の変化が同時に生じている.なお,血管の拡張及び収縮に起因する色の変化は非常に微弱であるため,爪から得られる測定波形は,前述の外

乱によって心拍成分がほとんど観測できない状態になる.本研究では,この問題を軽減するために,爪から得られた PPG に対して,音楽信号解析分野で発展した基本周波数推定法を適用し,ノイズに頑健な心拍推定が行えるかについて詳しく調査する.本論文では,以後 PPG 中の心拍に由来する波形成分を心拍信号と呼び,体動や筋電等の心拍以外に由来する波形成分をノイズと呼ぶ.

2.3 動作原理

本研究で用いる付け爪型センサの動作原理は Fig. 1.2 に示されている通りである. 付け爪型センサ本体には脈波計測回路, バッテリー, 無線送信機, 及び制御用マイコンが搭載されている. なお, ここで記載されている無線送信機について, Fig. 1.2 では有線で接続しているが, 将来的には無線通信で送信できるように改良予定である.

2.2 節でも説明したように, 爪先では血管の拡張及び収縮によって微細な色の変化が生じる. 本センサでは, これによって変動した光の反射率を利用することで脈波を計測している. 爪の表面から発された光は爪を透過し, 爪床に到達する. ここで, 爪床の反射光をフォトダイオード検出する. この反射光は, 増幅回路を通過した後に出力される. そして, 脈波計測回路にて出力された PPG は, 無線通信機により外部の受信機へ送られる. さらに, インターネット回線を通じて遠隔地にあるモニタリング端末へと情報を送ることが可能である.

また,光源には LED (LTSTC170TGKT)を使用し,波長は体動に伴う外乱の影響を受けにくい 530 nm と設定されている. なお,受光素子はフォトダイオード (TEMD6200FX01),増幅にはオペアンプ (OPA2330)を用いている. 計測回路は 11 mm×11 mm の基板に実装し,紫外線硬化樹脂によって包埋することで防水対策も施されている. これにより,水仕事及び入浴等の日常的な生活環境下での水没の危険性も回避している.

2.4 測定波形

文献 [3] で実施された 24 時間の測定実験で,付け爪型センサから得られた PPG 及び心臓付近に取り付けた電極から得られた ECG を Fig. 2.1 に示す.まず,Fig. 2.1 (a) で示す 24 時間全体の波形を見ると,大部分において波形がクリッピングしていることが分かる.これは,日常的な活動の動作時に生じる(心拍以外の)信号(本論文ではこれを外乱信号と考える)が加算されることによって,センサの測定可能範囲を超えてしまうことが原因である.しかしながら,Fig. 2.1 (b) 左列に示す睡眠時の拡大波形では,PPG 及び ECG のいずれの波形も心拍に由来すると思われる波形(本論文における心拍信号)が測定されている.ただし,ECG はスパイク状の目的信号が現れるのに対し,付け爪型センサで得られる PPG の目的信号は高周波成分をあまり含まない正弦波に近い波形となっている.Fig. 2.1 (b) 中央列の通常活動下や,Fig. 2.1 (c) 右列の激しい運動中では,運動の激しさに応じて外乱信号が増し,心拍に由来する信号の信号対雑音(signal-to-noise: SN)比が極端に低下する.

Japanese standard time	Scene
18:00-00:53	Moderate activity
00:53-07:00	Sleep
07:00-09:47	Moderate activity
09:47-11:07	Hard exercise
11:07-12:30	Moderate activity
12:30 - 13:56	Sleep
13:56-15:09	Hard exercise
15:09 - 18:00	Moderate activity

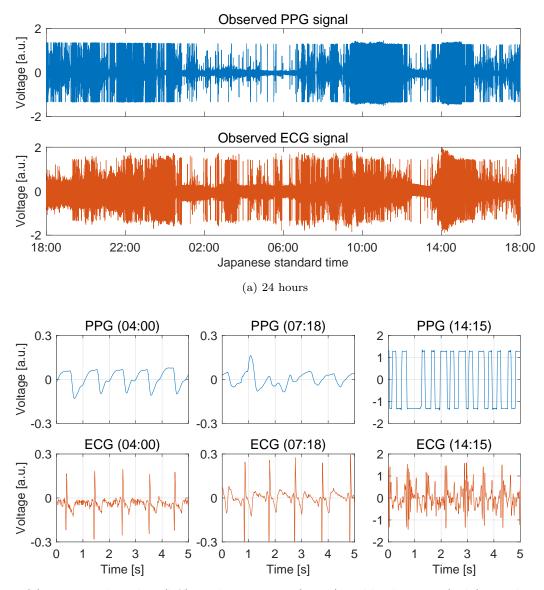
Table 2.1. Activity scenes in each time period (Data 1)

本論文では,文献 [1] で計測された $\mathrm{Fig.}\ 2.1$ の信号だけでなく,新たに計測した 2 日分の波形についても心拍推定の調査・解析を行う.これらの計測結果にそれぞれ「データ 1」、「データ 2」,及び「データ 3」と名付け,以後はこのデータ名を用いて計測波形を指し示す.各データの計測実験に関する詳細は以下の通りである.

データ 1 は 2018 年 5 月 2 日に測定した 24 時間のデータである.各信号のサンプリング周波数は 1 kHz である.このデータは,参考値として胸に取り付けた電極から得られた ECGの信号と,付け爪型センサから得られた PPG の信号の波形が,同じ A/D 変換器に入力されておらず,時間的な同期が取れていない.計測開始のタイミングは ECG と PPG で極力同時としたが,胸部の電極に接続される計測機器と付け爪型センサに接続される計測機器の間でサンプリング周波数はわずかにずれている可能性があり,24 時間という長時間の計測においては,波形後半の時間のずれが顕著となっている可能性がある.また,PPG においては,ダイナミックレンジを超えてクリッピングすることへの対策がされていない測定データである.データ 1 における各時刻での行動は 1 Table 1 の通りである.

データ 2 は 2021 年 5 月 3 日に測定した 24 時間のデータである.各信号のサンプリング周波数は $512~\mathrm{Hz}$ である.このデータは,参考値として胸に取り付けた電極から得られた ECG と付け爪型センサから得られた PPG が同じ A/D 変換器に入力されており,時間的な同期が取れている.PPG においては,ダイナミックレンジを超えてクリッピングすることへの対策として,受光センサに遮光テープを取り付けることにより,光の量の調整が施されている測定データである.データ 2 における各時刻での行動は Table 2.2 の通りである.Fig. 2.1 と同様に,データ 2 の波形を Fig. 2.2 に示す.

データ 3 は 2021 年 5 月 8 日に測定した 24 時間のデータである.各信号のサンプリング周波数は 1024 Hz である.ECG と PPG の時間同期についてはデータ 2 と同様であるが,データ 3 ではさらに LED にも遮光テープを貼り付け,より減光する処理が施されている.データ 3 における各時刻での行動は Table 2.3 の通りである.Fig. 2.1 と同様に,データ 3 の波形を Fig. 2.3 に示す.



(b) Five seconds in sleep (left), moderate activity (center), and hard exercise (right) periods
Fig. 2.1. PPG (top) and ECG (bottom) signals in Data 1 for (a) 24 hours and (b) five seconds.

2.5 本章のまとめ

本章では,付け爪型センサの概要について詳しく解説し,従来の胸部接触型の電極センサに対する利点を述べた.さらに,従来の胸部接触型の電極センサから得られる ECG の波形と比較したとき,付け爪型センサから得られる PPG では,外乱によって波形が大きく乱れるという問題点を述べた.次章では,本研究で提案する深層学習を用いての心拍推定の動機と手法について詳しく述べる.

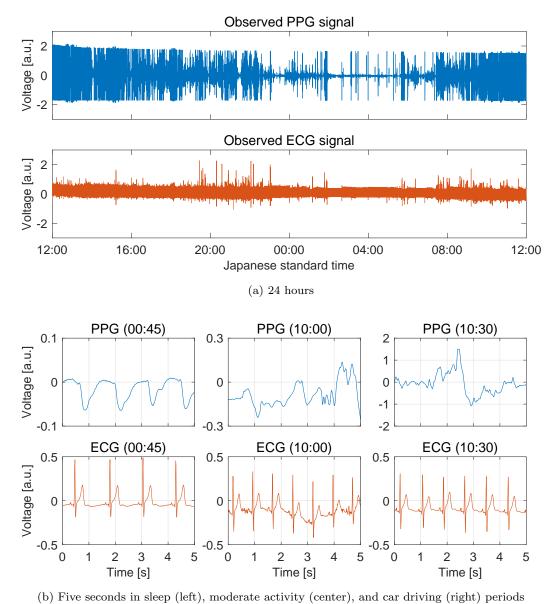


Fig. 2.2. PPG (top) and ECG (bottom) signals in Data 2 for (a) 24 hours and (b) five seconds.

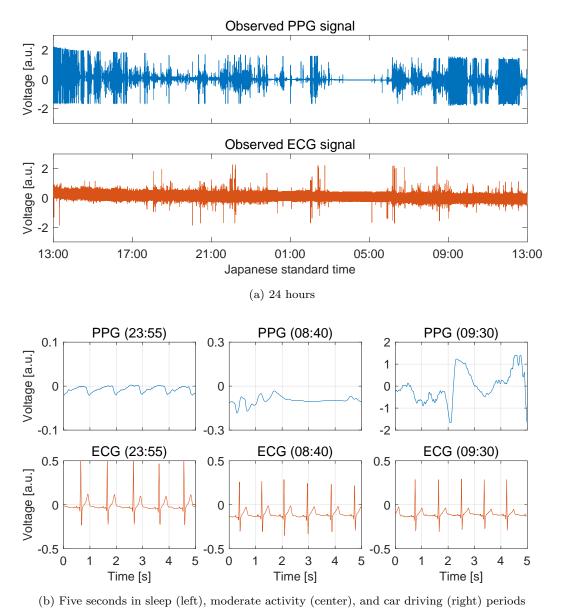


Fig. 2.3. PPG (top) and ECG (bottom) signals in Data 3 for (a) 24 hours and (b) five seconds.

Table 2.2. Activity scenes in each time period (Data 2)

Japanese standard time	Scene
12:00 - 12:50	Moderate activity
12:50-12:56	Car driving
12:56 - 13:00	Moderate activity
13:00 - 13:09	Car driving
13:09-14:50	Moderate activity
14:50-15:15	Walking
15:15-18:08	Moderate activity
18:08 - 18:30	Car driving
18:30-23:03	Moderate activity
23:03-07:30	Sleep
07:30 - 10:15	Moderate activity
10:15-10:41	Car driving
10:41 - 11:20	Moderate activity
11:20-11:35	Car driving
11:35-11:40	Moderate activity
11:40 - 12:09	Car driving

Table 2.3. Activity scenes in each time period (Data 3)

Japanese standard time	Scene
13:00 - 13:11	Moderate activity
13:11-13:34	Car driving
13:34-16:03	Moderate activity
16:03-16:17	Car driving
16:17-17:45	Moderate activity
17:45 - 18:05	Car driving
18:05-23:50	Moderate activity
23:50-06:10	Sleep
06:10-09:08	Moderate activity
09:08-09:53	Car driving
09:53-11:34	Moderate activity
11:34 - 11:48	Car driving
11:48 - 12:15	Moderate activity
12:15-12:32	Car driving
12:32-13:05	Moderate activity

第3章

基礎技術

3.1 まえがき

本章では,1章で述べた DNN に基づく心拍推定を構成する要素技術として,音響信号の時間周波数領域への変換である STFT について 3.2 節で説明する.また,本論文で構成する DNN モデルである MLP,RNN,及び BiRNN についてそれぞれ 3.3 節,3.4 節,及び 3.5 節で説明する.3.6 節で本章をまとめる.

3.2 STFT

STFT は Fig. 3.1 に示すように , 時間領域の信号からスペクトルを表現するための手法である . STFT において , 時間領域から時間周波数領域に変換するときの窓関数の長さ及びシフト 長をそれぞれ Q 及び τ とすると , 時間領域の信号 $z=[z(1),z(2),\cdots,z(t),\cdots,z(T)]^{\mathrm{T}}\in\mathbb{R}^T$ の j 番目の短時間区間 (時間フレーム) 信号は次式で表される .

$$\mathbf{z}^{(j)} = [z((j-1)\tau + 1), z((j-1)\tau + 2), \cdots, z((j-1)\tau + Q)]^{\mathrm{T}}$$
(3.1)

$$= \left[z^{(j)}(1), z^{(j)}(2), \cdots, z^{(j)}(q), \cdots, z^{(j)}(Q) \right]^{\mathrm{T}} \in \mathbb{R}^{Q}$$
(3.2)

ここで, $t=1,2,\cdots,T$, $j=1,2,\cdots,J$,及び $q=1,2,\cdots,Q$ は,それぞれ時間領域の信号の離散インデクス,時間フレーム,及び時間フレーム内のサンプルのインデクスであり, \cdot^{T} は転置を表す.また,時間フレーム J は次式で与えられる.

$$J = \frac{T}{\tau} \tag{3.3}$$

ただし,J は自然数となる必要があるため,信号長 T がシフト長 au で割り切れるように信号の末尾を零埋めする.この時,信号 z の複素スペクトログラム D を次式で表す.

$$\mathbf{Z} = \operatorname{STFT}_{\omega}(\mathbf{z}) \in \mathbb{C}^{I \times J}$$
 (3.4)

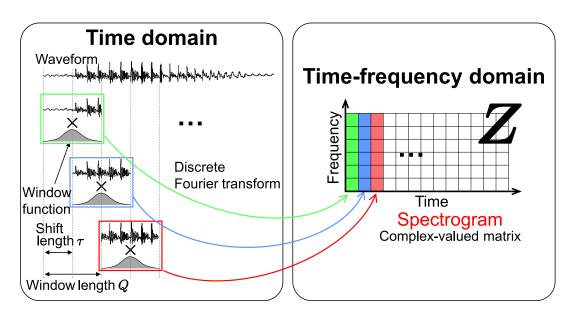


Fig. 3.1. Mechanism of STFT.

ここで, $\omega=[\omega(1),\omega(2),\cdots,\omega(q),\cdots,\omega(Q)]^{\mathrm{T}}\in\mathbb{R}^Q$ は窓関数を表す.この時,複素スペクトログラム Z の (i,j) 成分 z_{ij} は次式で表される.

$$z_{ij} = \sum_{q=1}^{Q} \omega(q) z^{(j)}(q) \exp \frac{-2\pi\iota(q-1)(i-1)}{F}$$
(3.5)

ここで F は $\lfloor \frac{F}{2} \rfloor + 1 = I$ を満たす整数 , ($\lfloor \cdot \rfloor$ は床関数) , $i=1,2,\cdots,I$ は周波数ビンのインデクス , ι は虚数単位を表している.また , 時間周波数成分の大きさのみを取り扱うことも少なくない.この場合は , 複素スペクトログラム Z の各要素に関して絶対値をとった振幅スペクトログラム $|Z| \in \mathbb{R}^{I \times J}_{\geq 0}$ や , 絶対値の 2 乗を取ったパワースペクトログラム $|Z|^{\cdot 2} \in \mathbb{R}^{I \times J}_{\geq 0}$ を処理の対象とする.ここで , 行列に対しての絶対値記号及びドット付指数乗は行列の要素毎の絶対値及び指数乗を表す.

3.3 MLP に基づく DNN

DNN の基本形である MLP に基づく DNN は , Fig. 3.2 のようにニューラルネットワークを多層に重ねたネットワーク構造で構成されており , ニューラルネットワークの基本単位である「層」が重要となる . まず , ニューラルネットワークが MLP に基づく DNN においてどのような役割を果たしているかを説明する .

ニューラルネットワークとは,出力データが入力データの重みづけと非線形写像で表現できるようなネットワーク構造である.中でも,入力データを構成する個々の要素がすべての出力データの要素と結びついているものを全結合層(dense layer)と呼び,本論文で用いている MLP はこの全結合層を用いて構成している.今,入力データを $x=[x_1,x_2,\cdots,x_N]^{\mathrm{T}}\in\mathbb{R}^N$ とする.全結合層の構造は Fig. 3.3 のようになっており,全結合層の各出力 $y_{n'}$ は次式のよう

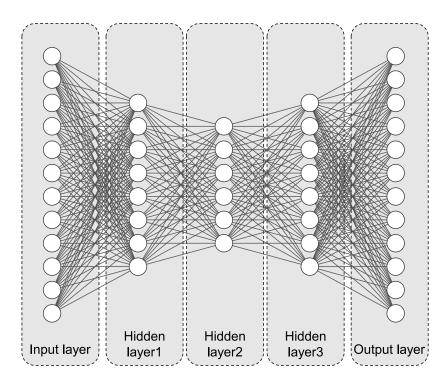


Fig. 3.2. Multilayer neural network.

に表すことができる.

$$y_{n'} = \phi \left(\sum_{n=1}^{N} w_{nn'} x_n + b_n \right)$$
 (3.6)

ここで, $n=1,2,\cdots,N$ は入力データの要素インデクス, $n'=1,2,\cdots,N'$ は出力データの要素インデクス, $w_{nn'}$ は重み係数を表す.また ϕ は活性化関数と呼ばれ,入力データの総和をどのように出力に反映させるかを決定する役割を持つ.これは次の層に渡す値を整えるような非線形関数であり, ${
m Fig.}~3.4$ のように様々な活性化関数が定義されている.

式 (3.6) のニューラルネットワーク 1 層を基本単位として多層に重ねたものを DNN と呼ぶ、本節では,例として隠れ層が 3 層で構成される DNN を取り扱う.Fig.~3.2 に示すように,データが最初に入力される層を入力層,入力されたデータが伝達されていく層を隠れ層(中間層),最終的に値が出力される層を出力層と呼ぶ.DNN の学習は,Fig.~3.2 の出力層から得られる予測結果とラベル(予測結果の正解値)の間でなんらかの誤差関数を定義し,これを最小化するように DNN 内のすべての重み係数を求めることで実現できる.また,そのような最適化は誤差逆伝播と呼ばれる方法により効率的に実行される.

3.4 RNN に基づく DNN

 MLP は信号の持つ構造や系列方向等をすべて無視して 1 本のベクトルとして入力データを構築し,入力層に与えるモデルである.一方, RNN に基づく DNN は,時刻毎の観測値を持

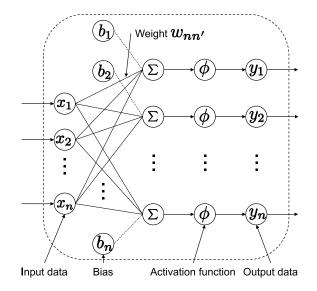


Fig. 3.3. Structure of dense (fully connected) layer.

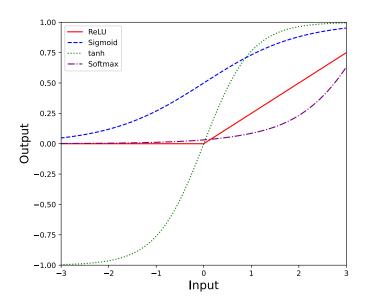


Fig. 3.4. Popular activation functions used in DNN.

つ信号のように,何らかの系列となっている信号を対象とする.この時,系列方向の再帰性をネットワーク構造として陽に定義することで,信号の持つ系列方向の関連性を MLP よりも効率的に保ったまま学習することができる.本論文のデータも時系列データであるため,RNNを利用することは有用であると考えられる.

RNN では, ${
m Fig.~3.5}$ のようにある系列の隠れ層からの出力が,次の系列における入力に用いられる構造を持っている.この構造を系列方向に展開して考えた際に,時刻 t における中間層の出力は次式で得られる.

$$\boldsymbol{h}_{t} = \varphi\left(\boldsymbol{x}_{t}, \boldsymbol{h}_{t-1}, \boldsymbol{W}^{\boldsymbol{x}\boldsymbol{h}}, \boldsymbol{W}^{\boldsymbol{h}\boldsymbol{h}}\right)$$
(3.7)

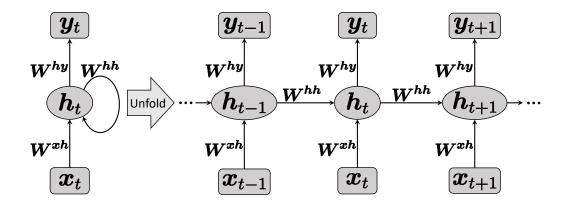


Fig. 3.5. Simplified diagram of RNN structure.

ここで, W^{xh} 及び W^{hh} はそれぞれ入力層から隠れ層に伝搬する際の重み係数及び隠れ層の出力が次の時系列の隠れ層に伝搬する際の重み係数となっている. φ は活性化関数を含むネットワーク構造を表す非線形関数である.また,出力層の値は次式で得られる.

$$\mathbf{y}_t = \varphi'\left(\mathbf{h}_t, \mathbf{W}^{h\mathbf{y}}\right) \tag{3.8}$$

ここで, W^{hy} は隠れ層から出力層に伝搬する際の重み係数であり, φ' は φ と同様に活性化関数を含むネットワーク構造を表す非線形関数である.

このようなネットワーク構造のもと,予測結果とラベルの間の誤差関数を最小化するようにすべての層の重み係数ベクトルを学習する.RNN における最適化の方法は MLP と同様である.

3.4.1 LSTM

LSTM とは,RNN の隠れ層に,LSTM ユニット [19] を用いた構造を持つ DNN である. Fig. 3.6 に LSTM ユニットの構造を示す.Fig. 3.6 の各変数の計算は次式となる.

$$\mathbf{f}_t = \sigma(\mathbf{W}^f \mathbf{x}_t + \mathbf{R}^f \mathbf{h}_{t-1} + \mathbf{b}^f) \tag{3.9}$$

$$\tilde{\boldsymbol{i}}_t = \sigma(\boldsymbol{W}^{\tilde{\boldsymbol{i}}} \boldsymbol{x}_t + \boldsymbol{R}^{\tilde{\boldsymbol{i}}} \boldsymbol{h}_{t-1} + \boldsymbol{b}^{\tilde{\boldsymbol{i}}})$$
(3.10)

$$\tilde{\boldsymbol{c}}_t = \tanh(\boldsymbol{W}^{\tilde{\boldsymbol{c}}} \boldsymbol{x}_t + \boldsymbol{R}^{\tilde{\boldsymbol{c}}} \boldsymbol{h}_{t-1} + \boldsymbol{b}^{\tilde{\boldsymbol{c}}})$$
(3.11)

$$\boldsymbol{c}_t = \boldsymbol{f}_t \circ \boldsymbol{c}_{t-1} + \tilde{\boldsymbol{i}}_t \circ \tilde{\boldsymbol{c}}_t \tag{3.12}$$

$$o_t = \sigma(\mathbf{W}^o \mathbf{x}_t + \mathbf{R}^o \mathbf{h}_{t-1} + \mathbf{b}^o) \tag{3.13}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \circ \tanh(\boldsymbol{c}_t) \tag{3.14}$$

ここで, $\sigma(\cdot)$ はシグモイド関数, $anh(\cdot)$ は双曲線正接関数, \circ はベクトルの要素毎の積をそれぞれ示す. W^f , $W^{ ilde t}$,及び W^o は時間 t における入力ベクトル x_t に対する重み係数

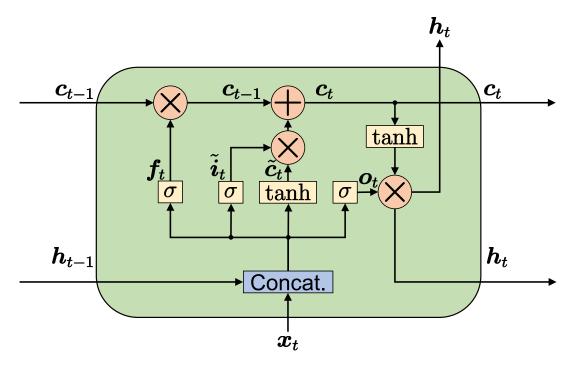


Fig. 3.6. Structure of LSTM unit.

行列, R^f , $R^{\tilde{i}}$, $R^{\tilde{c}}$,及び R^o は時間 t-1 における出力ベクトル h_{t-1} に対する重み係数行列, b^f , $b^{\tilde{i}}$, $b^{\tilde{c}}$,及び b^o はそれぞれの係数に対するバイアスベクトルを示す.また, f_t , \tilde{i}_t ,及び o_t はそれぞれ忘却ゲート,入力ゲート,及び出力ゲートと呼ばれている.忘却ゲートは c_{t-1} からどの要素を保持するか決定するベクトルである.入力ゲートは x_t 及び h_{t-1} からどの要素を保持するか決定するベクトルである.出力ゲートは h_t を求めるためのベクトルである.Fig. 3.6 において,出力は h_t であり,次の LSTM ユニットに入力されるのは h_t 及び c_t である.

3.4.2 GRU

 ${
m GRU}$ は LSTM と異なり,単一のゲートユニットが忘却の要因と状態ユニットの更新の決定を同時に制御する. ${
m GRU}$ の内部を ${
m Fig.}~3.7$ に示す. ${
m Fig.}~3.7$ の各変数の計算は次式となる.

$$r_t = \sigma(\mathbf{W}^r \mathbf{x}_t + \mathbf{R}^r \mathbf{h}_{t-1} + \mathbf{b}^r) \tag{3.15}$$

$$\tilde{\boldsymbol{h}}_{t} = \tanh(\boldsymbol{W}^{\tilde{\boldsymbol{h}}} \boldsymbol{x}_{t} + \boldsymbol{R}^{\tilde{\boldsymbol{h}}} (\boldsymbol{r}_{t} \circ \boldsymbol{h}_{t-1}) + \boldsymbol{b}^{\tilde{\boldsymbol{h}}})$$
(3.16)

$$\tilde{\mathbf{z}}_t = \sigma(\mathbf{W}^{\tilde{\mathbf{z}}}\tilde{\mathbf{z}}_t + \mathbf{R}^{\tilde{\mathbf{z}}}\mathbf{h}_{t-1} + \mathbf{b}^{\tilde{\mathbf{z}}}) \tag{3.17}$$

$$\boldsymbol{h}_{t} = (1 - \tilde{\boldsymbol{z}}_{t}) \circ \boldsymbol{h}_{t-1} + \tilde{\boldsymbol{z}}_{t} \circ \tilde{\boldsymbol{h}}$$

$$(3.18)$$

ここで, $m{W}^{ ilde{i}}$, $m{W}^r$,及び $m{W}^{ ilde{h}}$ は時間 t における入力ベクトル $m{x}_t$ に対する重み係数行列, $m{R}^{ ilde{z}}$, $m{R}^r$, $m{R}^{ ilde{h}}$,及び $m{R}^o$ は時間 t-1 における出力ベクトル $m{h}_{t-1}$ に対する重み係数行列, $m{b}^{ ilde{z}}$, $m{b}^r$, $m{b}^{ ilde{h}}$,及び $m{b}^o$ はそれぞれの係数に対するバイアスベクトルを示す. $m{r}_t$ 及び $m{z}_t$ はそれぞれリ

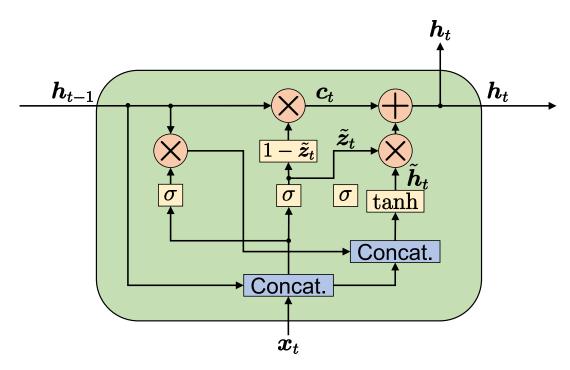


Fig. 3.7. Structure of GRU.

セットゲート及び更新ゲートと呼ばれている.リセットゲートは $m{h}_t$ からどの要素を保持するか決定するベクトルである.更新ゲートは $m{x}_t$ 及び $m{h}_{t-1}$ からどの要素を保持するか決定するベクトルである.

3.5 BiRNN に基づく DNN

BiRNN に基づく DNN は , Fig. 3.8 に示すように過去から未来の方向(以後順方向とする)に処理する RNN と未来から過去の方向(以後逆方向とする)に処理する RNN を組み合わせた DNN である . 順方向の RNN の中間層の出力は次式で得られる .

$$h_t = \varphi\left(x_t, h_{t-1}, W_{\text{forward}}^{xh} W_{\text{forward}}^{hh}\right)$$
 (3.19)

ここで, $W^{xh}_{
m forward}$ 及び $W^{hh}_{
m forward}$ はそれぞれ順方向の RNN における入力層から隠れ層に伝搬する際の重み係数及び隠れ層の出力が次の時系列の隠れ層に伝搬する際の重み係数となっている.一方,逆方向の RNN の中間層の出力は,中間層の出力を新たに g_t と定義すると次式で得られる.

$$g_t = \varphi\left(x_t, g_{t+1}, W_{\text{backward}}^{xg}, W_{\text{backward}}^{gg}\right)$$
 (3.20)

ここで, $W^{xg}_{
m backward}$ 及び $W^{gg}_{
m backward}$ はそれぞれ逆方向の m RNN における入力層から隠れ層に 伝搬する際の重み係数及び隠れ層の出力が次の時系列の隠れ層に伝搬する際の重み係数となっ

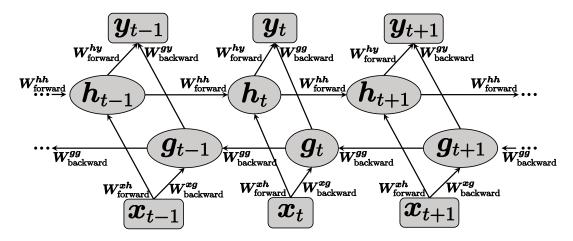


Fig. 3.8. Simplified diagram of BiRNN structure.

ている.また,出力層の値は次式で得られる.

$$\boldsymbol{y}^{(t)} = \varphi'\left(\boldsymbol{h}^{(t)}, \boldsymbol{g}^{(t)}, \boldsymbol{W}_{\text{forward}}^{\boldsymbol{h}\boldsymbol{y}}, \boldsymbol{W}_{\text{backward}}^{\boldsymbol{g}\boldsymbol{y}}\right)$$
 (3.21)

このように,BiRNN は系列データに対して順方向と逆方向の RNN をそれぞれ独立に適用し,その出力を式 (3.21) の形でその系列の出力に写像させる構造となっている.この順方向と逆方向の RNN に LSTM ユニットを使えば BiLSTM となり,GRU を使えば BiGRU となる.

3.6 本章のまとめ

本章では,1章で述べた DNN に基づく心拍推定を構成する要素技術について説明した.3.2節では,音響信号処理でよく用いられる STFT について説明した.3.3 節では,最も基本的な DNN である MLP に基づく DNN について説明した.3.4 節では,再帰性を考慮した学習が可能である DNN である RNN に基づく DNN について説明した.また,中間層として用いられる LSTM 及び GRU についても説明した.3.5 節では,2 つの RNN を同時に用いる DNN である BiRNN 型 DNN について説明した.次章では,本論文における提案手法として,本章で説明した要素技術を用いた心拍推定法について説明する.

第4章

提案手法

4.1 まえがき

本章では,1章で説明した本論文の提案手法である DNN に基づく心拍推定の詳細を説明する.4.2節では,本研究で DNN に基づく教師あり心拍推定を実現することの動機と,ラベリングアプリケーション開発の必要性について述べる.4.3 節では,DNN の学習に必要となるラベル付き学習データを準備するためのラベリングアプリケーションの開発について説明する.4.4 節では,測定した PPG を DNN に入力するにあたって施した前処理について説明する.4.5 節では,各 DNN の入力について説明する.4.5 節では,各 DNN の入力について説明する.4.5 節では,各 DNN の出力及び損失関数の測り方について説明する.4.7 節では,本章をまとめる.

4.2 提案手法の動機

医療行為として一般的に測定される心拍信号(ECG)は,被験者が仰臥位等で安静にした状態で計測される.そのため,ノイズの少ない波形が観測信号として得られ,ゼロクロス法等の簡易的な手法で心拍値を得ることが可能となる.しかしながら,本研究で用いる付け爪型センサ[3]では2.2節で説明したように,外乱によるノイズが多く混入し,特に日常生活時においてはSN 比が著しく低下する.そのため,ゼロクロス法では心拍の推定はほとんど困難である.

1.2 節でも述べたように,付け爪型センサで得られる PPG から心拍を推定する方法として, 文献 [12,13] では音楽信号処理における 4 つの基本周波数推定法が適用された.その結論として,睡眠時のような安静にしている状態では心拍を推定できるが,日中の活動時の心拍推定はほとんど困難であった.しかしながら,データ 1 からデータ 3 の PPG に STFT を適用して得られるパワースペクトログラム(それぞれ Figs. 4.1–4.3.但し,4.4 節で述べるハイパスフィルタ適用後である)を見ると,日常生活においても多くの時間において,SN 比は悪いながらも心拍の基本周波数が確認できる個所が存在する(但し,データ 3 は PPG のパワーそのものが小さく,ほとんど確認できない).従来の教師無し手法では,これらの PPG から心拍値を求めることは難しかったが,学習データを最大限活用する教師あり手法では実現できる可能

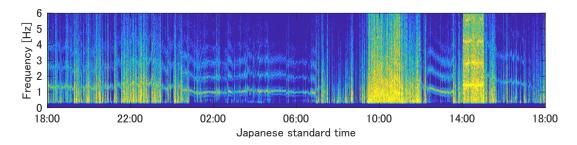


Fig. 4.1. Power spectrogram of Data 1 (PPG) after applying high-pass filter.

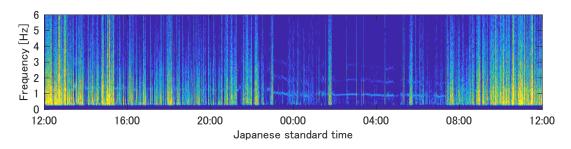


Fig. 4.2. Power spectrogram of Data 2 (PPG) after applying high-pass filter.

性が高いと考えられる.そこで本研究では,日中の活動時の心拍推定を目標として DNN に基 づく教師あり心拍推定法を提案する.この時,各時刻において正解の心拍値が与えられた(ラ ベル付きの) PPG が必要となる .2 章で述べたように, データ1 からデータ3 はすべて接触 型心拍計を用いた ECG も同時に取得しており,この ECG から心拍値を求めることでラベル を与えることができると予想されるが,文献[12,13]で示した通り, ECG に対してもあまり 有効な心拍推定は現状できておらず,信頼性の高いラベルを付与することが難しい.さらに本 研究で目指す心拍推定では, PPG において心拍信号が完全に見えなくなっている期間におい ては「計測不可」という結果を出力するようなシステムを学習したい、以上の目的を満たす教 師あり心拍推定法を提案するために,本研究ではまず,ユーザが手動で(目視で)ラベルを簡 便に付与することのできるラベリングアプリケーションを開発する.これは , ${
m Figs.~4.1-4.3}$ に 示したように,スペクトログラムからある程度心拍信号の基本周波数が確認できることに着想 を得たものである.このラベリングアプリケーションは,ユーザがある程度の確信度を持った 期間にのみラベルを付与できるよう設計されており,心拍の基本周波数が確認できない所はラ ベルを付与しないで良い、さらに、このラベリングアプリケーションを利用すれば、今後の学 習データの収集において ECG を計測する必要が無くなるため,より手軽に PPG の学習デー タを得ることが可能となる.このようにラベルを付与した PPG を学習データとして DNN を 学習し,その精度について次章で確認する.

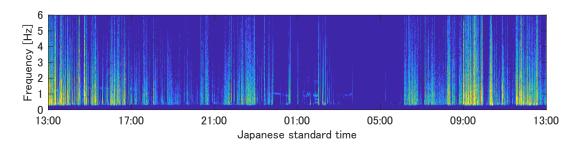


Fig. 4.3. Power spectrogram of Data 3 (PPG) after applying high-pass filter.

4.3 ラベリングアプリケーションの開発

Figs. 4.1—4.3 に示す PPG のパワースペクトログラムでは,0.7—5.0 Hz のあたりに横筋状の縞模様が確認できる.これは,PPG に含まれる心拍信号の周波数成分であり,最も低い周波数の横筋が基本周波数(すなわち心拍)に対応しており,その整数倍が高調波成分である.人間の心拍数は基本的に 36—210 bpm (0.6—3.5 Hz) となっていることからも,この横筋状の縞模様が心拍信号の成分であることは妥当と考えられる.特に,睡眠時(例えば,Fig. 4.1 の 01:00—07:00 等)は良好な SN 比で心拍信号の成分が確認できる.本節では,このように目視で確認できる心拍信号成分の基本周波数をユーザがマウスやスタイラスペン等のデバイスでなぞることで,直感的に心拍値のラベル付けができるラベリングアプリケーションについて説明する.

本研究で実際に開発したラベリングアプリケーションの画面を ${
m Fig.}~4.4$ に示す.また,このアプリケーションの機能の概要を以下に示す.なお,このラベリングアプリケーションは ${
m MATLAB}$ の ${
m App}$ ${
m Designer}$ を用いて開発した.

- アプリケーションのファイル入力画面で CSV 形式の PPG を指定すると,ファイル内の信号値が読み込まれ,STFT が実行されて,パワースペクトラムが描画される.
- マウスやスタイラスペン等のデバイスを用いてパワースペクトログラム上をドラッグすると,赤色の線としてラベルの曲線が表示される(FIg. 4.5).この曲線の縦軸のインデクス(画面上では周波数ビン番号)が各時間フレーム(パワースペクトログラムの各列)の心拍値ラベルとなる.この操作は繰り返し行うことができ,複数の場所にラベルとなる赤色の曲線を付与できる.
- ●「1つ戻る」ボタンを押すと最後に描画したラベル曲線を削除する.
- ●「YLimMax」及び「YLimMin」の数値バーを調整して縦軸の描画範囲を指定する.
- ●「XWidth」及び「XPos」の数値バーを調整して横軸の描画範囲と描画位置を変更できる.
- ●「CRangeMax」及び「CRangeMin」の数値バーを調整してカラーマップを調整で

Fig. 4.4. User interface of proposed labeling application.

きる.

● ラベル付けが完了した後は、「確定」ボタンをクリックすることで、各時間フレームにおけるラベル曲線の縦軸のインデクスをまとめた行ベクトルが CSV 形式のファイルとして出力される.この時、ラベルが振られなかった(ラベル曲線が描かれなかった)時間フレームには 0 という数値がラベルとして付与される.また、1 つの時間フレームに複数本のラベル曲線が描かれた場合は、最後に描いたラベル曲線の縦軸のインデクスが優先されて出力される.

実際にこのラベリングアプリケーションを用いて作成したラベルを Figs.~4.6-4.8 に示す.但し,縦軸の値は周波数ビン番号から周波数値に変更して表示している.いずれのデータのラベリングにおいても,SN 比が極端に悪く心拍信号の基本周波数が判別できない期間についてはラベル曲線を付与しなかった.そのような時間フレームは0 というラベルが付けられている.

4.4 PPG の前処理

観測した PPG を $\tilde{s}_{\text{raw}}=[\tilde{s}(1),\tilde{s}(2),\cdots,\tilde{s}(t),\cdots,\tilde{s}(T)]\in\mathbb{R}^T$ とする.PPG として測定される電圧波形には直流成分が含まれている(但し,Figs. 2.1–2.3 の波形は波形全体の平均値を減算して表示している).信号の基本周波数を解析する上では,この直流成分を事前に除去し

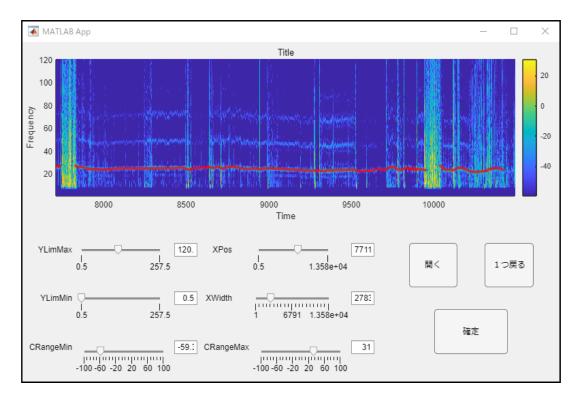


Fig. 4.5. Screen capture of labeling process.

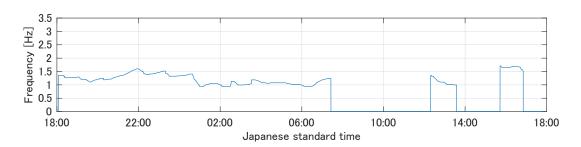


Fig. 4.6. Label of Data 1 (PPG).

ておくことが望ましいため,前処理としてハイパスフィルタを適用する.本研究では, $\tilde{s}_{\rm raw}$ に対して Butterworth IIR ディジタルハイパスフィルタを用いて,直流成分を除去した信号 \tilde{s} に変換する.この時,カットオフ周波数は $0.3~{\rm Hz}$ 及びフィルタ次数は $30~{\rm C}$ 次に設定している.このフィルタの振幅応答と位相応答を ${\rm Fig.}~4.9~{\rm C}$ に示す.

また,本論文では,測定した PPG のサンプリング周波数を観測時の周波数(データ 1 は $1000~{\rm Hz}$,データ 2 は $1024~{\rm Hz}$,データ 3 は $512~{\rm Hz}$)から $20~{\rm Hz}$ にダウンサンプリングする.ダウンサンプリンングは以下の手順で行っており,これはディジタル信号処理における一般的なダウンサンプリング法である.

1. エイリアシングによる歪みを避けるため,ダウンサンプリング前の信号に対して FIR ローパスフィルタ(フィルタ次数は 1000 点,カットオフ周波数はダウンサンプリング

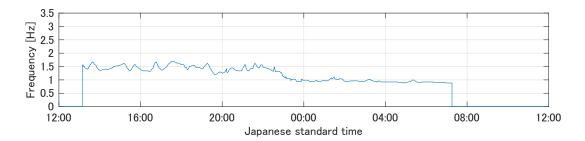


Fig. 4.7. Label of Data 2 (PPG).

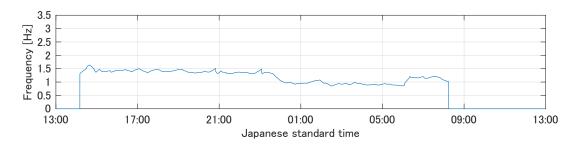


Fig. 4.8. Label of Data 3 (PPG).

後のナイキスト周波数)を適用する

- 2. FIR ローパスフィルタで生じた信号の群遅延 (波形全体の時間遅れ)を補償する
- 3. ダウンサンプリング後の離散時間の各信号値を線形内挿(線形補間)により求める

ダウンサンプリングした信号 \tilde{s} のパワースペクトログラムは次式で得られる.

$$S = |\text{STFT}_{\omega}(\tilde{s})|^{\cdot 2} \in \mathbb{R}^{I \times J}$$
(4.1)

$$= [\mathbf{s}_1 \ \mathbf{s}_2 \ \cdots \ \mathbf{s}_j \ \cdots \ \mathbf{s}_J] \tag{4.2}$$

また, s_i はパワースペクトログラム S の各列ベクトルを示し,次式で定義する.

$$\mathbf{s}_j = [s_{1j}, s_{2j}, \cdots, s_{ij}, \cdots, s_{Ij}]^{\mathrm{T}} \in \mathbb{R}^I$$
(4.3)

ここで, s_{ij} はパワースペクトログラム S の (i,j) 番目の要素を示す.データ 1,データ 2,及 びデータ3のパワースペクトログラムS はそれぞれFigs. 4.1-4.3の通りである.

また,DNNの入力として用いるため,パワースペクトログラムに対して正規化を行う.正 規化を適用したパワースペクトログラムは次式で得られる.

$$\overline{S} = \left[\frac{s_1}{\max(s_1)}, \frac{s_2}{\max(s_2)}, \cdots, \frac{s_J}{\max(s_J)} \right] \in [0, 1]^{I \times J}$$
(4.4)

ここで, $\max(\cdot)$ は入力されたベクトルの最大値を示す.

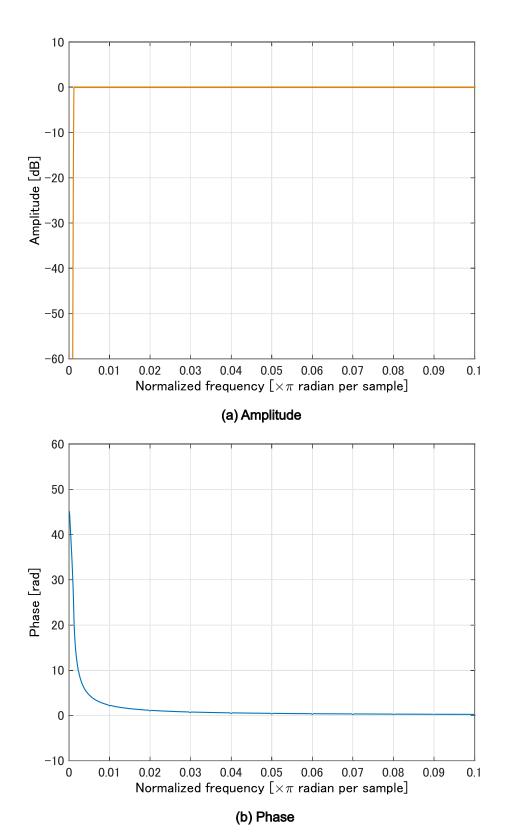


Fig. 4.9. Characteristic of high-pass filter used in PPG preprocessing: (a) amplitude and (b) phase responses.

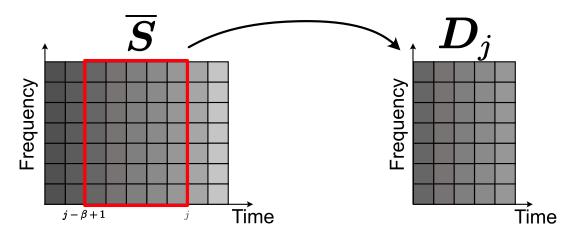


Fig. 4.10. Extraction of local-time-frame power spectrogram.

4.5 DNN モデルの入力

Fig. 4.10 に示すように,現在の時間フレーム j を含む過去 β 個の時間フレームを部分的なパワースペクトログラムとして抽出し,これを DNN の入力データとする.この処理は次式で表される.

$$\mathbf{D}_{i} = \left[\overline{\mathbf{s}}_{i-\beta+1} \ \overline{\mathbf{s}}_{i-\beta+2} \ \cdots \ \overline{\mathbf{s}}_{i} \right] \in \left[0, 1 \right]^{I \times \beta} \tag{4.5}$$

ここで, β (0 以上の整数)は離散時間フレームをどの程度 DNN に入力するかを決めるパラメータである.また, \overline{s}_j は \overline{S} の j 番目の列ベクトルである.本論文で取り扱う 5 つの DNN モデルにおいて,MLP に基づく DNN は入力をベクトルにする必要がある.この処理は次式で実現される.

$$\boldsymbol{d}_j = \operatorname{vec}(\boldsymbol{D}_j) \in [0, 1]^{I\beta} \tag{4.6}$$

ここで, $\text{vec}(\cdot)$ は,入力された行列をベクトルに整形する処理を表す.また,RNN に基づく DNN 及び BiRNN に基づく DNN は,時間フレーム方向の系列を保ったまま入力するため式 (4.5) で与えられる D_i そのものとなる.

4.6 DNN モデルの出力と学習

4.5 節で述べたように , MLP に基づく DNN の予測は次式で得られる .

$$u_j = \text{DNN}(\boldsymbol{d}_j) \in [0, 3.5] \tag{4.7}$$

また, RNN 及び BiRNN の出力はいずれも次式で得られる.

$$u_i = \text{DNN}(\boldsymbol{D}_i) \in [0, 3.5] \tag{4.8}$$

ここで, u_j は離散時間フレーム j に対する心拍の予測値(単位は Hz)と定義する.従って, u_j はラベリングアプリケーションで作成した正解値に近づくように DNN を学習する.但し,人間の心拍数は基本的に $3.5~\mathrm{Hz}$ ($210~\mathrm{bpm}$)より小さいため, DNN の予測値の最大値は $3.5~\mathrm{Hz}$ となるようにネットワーク内で制限している.その具体的な方法は $5.3~\mathrm{mod}$ 節で述べる.

DNN の学習は,何らかの損失関数を定義し,その値を最小化するパラメータを誤差逆伝搬により推定する処理である.心拍を推定する問題は回帰問題であるため,本論文では一般的に回帰モデルを作成する際に用いられる損失関数である平均二乗誤差を使用する.ラベリングアプリケーションで作成されたラベルを $m v=[v_1,v_2,\cdots,v_j,\cdots,v_J]\in\mathbb{R}^J$ とすると損失関数は以下で表される.

$$\mathcal{L}(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{J} \sum_{j=1}^{J} (u_j - v_j)^2$$
(4.9)

4.7 本章のまとめ

本章では,1章で説明した DNN に基づく心拍推定についての詳細と,この実現のために必要な正解値の作成方法について説明した.4.2 節では,本研究を行うにあたっての動機について述べた.4.3 節では,DNN の学習に必要となる正解値を付与するために開発したアプリケーションの機能について説明し,外観を示した.4.4 節では,測定した PPG を DNN に入力するにあたって施した IIR ディジタルフィルタ,ダウンサンプリング,STFT,及び正規化について説明した.4.5 節では,各 DNN の入力について説明した.4.6 節では,各 DNN の出力及び損失関数の測り方について説明した.50 節では,51 の出力及び損失関数の測り方について説明した.52 節では,53 節で作成したラベルを用いて実際に各 DNN の学習を行い,心拍推定の予測精度の評価を行う.

第5章

心拍の予測実験

5.1 まえがき

本章では,DNN に基づく心拍推定法の実験の詳細及び結果について説明する.5.2 節では,DNN に基づく心拍推定法によって得られた予測値に対する評価指標について説明する.5.3 節では,DNN の学習モデルの構造,データの分割方法,及び実験条件について説明する.5.4 節では,5.3 節で説明した実験条件に基づき,MLP,LSTM,GRU,BiLSTM,DND 及び DND に対してそれぞれ実験を行い,テストデータに対する心拍の予測結果を示す.5.5 節では,本章をまとめる.

5.2 評価指標

本実験では DNN によって得られた心拍の予測の精度を評価する.評価指標として,絶対値 誤差(mean absolute error: MAE),二乗平均平方根誤差(root mean squared error: RMSE),及びピアソン相関(Pearson correlation)を用いる.心拍の正解値は 4.3 節で作成した値を用いる.各評価指標について以下に示す.

MAE は予測値と正解値の絶対誤差の平均値を評価値とする指標である.MAE は次式で表される.

$$MAE = \frac{1}{J} \sum_{j=1}^{J} |u_j - v_j|$$
 (5.1)

RMSE は予測値と正解値の差を二乗し,この平均の平方根を評価値とする指標である.RMSE は次式で表される.

RMSE =
$$\frac{1}{J} \sum_{i=1}^{J} \sqrt{(u_j - v_j)^2}$$
 (5.2)

ピアソン相関は予測値と正解値の線形の相関を表す.ピアソン相関の結果は [-1,1] の範囲で表され,1 に近いほど正の相関が大きく,-1 に近いほど負の相関が大きい.ピアソン相関は

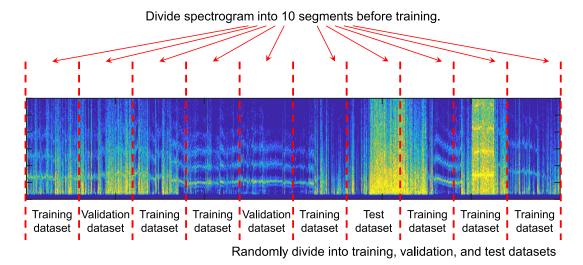


Fig. 5.1. Method of division data.

次式で表される.

$$P_{\text{Corr}} = \frac{\sum_{j=1}^{J} (u_j - \overline{u}) (v_j - \overline{v})}{\sqrt{\sum_{j=1}^{J} (u_j - \overline{u})^2} \sqrt{\sum_{j=1}^{J} (v_j - \overline{v})^2}}$$
(5.3)

ここで \overline{u} 及び \overline{u} はそれぞれ予測値の平均及び正解値の平均である.

5.3 実験条件

本実験ではデータ 1 からデータ 3 を用いて心拍推定を行う.しかしながら,これらはすべて 測定の際の減光処理が異なる.そのため,学習に用いるデータを測定条件が同じである単一の データのみの実験と,測定条件が異なるデータを含むすべてのデータを用いた実験の 2 種類の 実験を行う.

また,DNN の学習の際の学習データ,検証データ,及びテストデータの割合は 7:2:1 とし,データの重複が生じないようにする.データが重複しないようにするため, $Fig.\ 5.1$ に示すようにデータをあらかじめ 100 分割してから 7:2:1 に分ける.

本実験で用いた MLP , RNN , 及び BiRNN の構造を Figs. 5.2–5.4 に示す . MLP の隠れ層は 4 層とし隠れ層の次元はすべて同じ 100 次元とする . また , BiRNN において順方向及び逆方向の 2 本の出力ベクトルはそのまま結合 (concatenation) し , 1 層の全結合層を通して心拍の 1 次元に圧縮する . いずれの DNN においても , 最終層の活性化関数に sigmoid 関数を用いてから 3.5 倍することで , 予測値の範囲を [0,3.5] に制限している . その他の実験条件を Table 5.1 に示す . 学習のエポック数や入力する時間フレーム数はモデルによって共通とした .

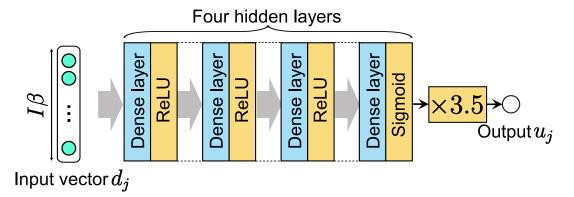


Fig. 5.2. MLP architecture.

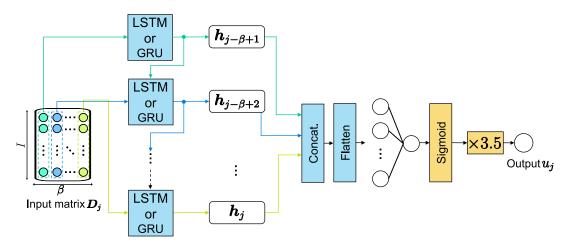


Fig. 5.3. RNN architecture.

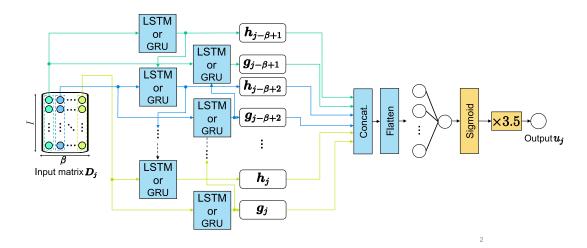


Fig. 5.4. BiRNN architecture.

5.4 実験結果

本節では,5.3節で説明した条件に基づく実験の結果の内,学習にデータ2のみを用いた実験及び学習にデータ1からデータ3のすべてを用いた実験の実験結果をそれぞれ5.4.1項及び

Window function in STFT	Blackman window
Window length in STFT	512 points (25.6 sec)
Shift length	128 points (6.4 sec)
Number of epochs	100
Number of time frames after division	134
Optimizer	Adam
Learning rate	10^{-5}

Table 5.1. Experimental conditions

5.4.2 項に示し,評価を行う.学習にデータ1 のみを用いた実験及び学習にデータ3 のみを用いた実験の実験結果については本文では割愛し,それぞれ付録A 及びB に掲載する.

5.4.1 データ2のみを用いた学習と評価

本項では,データ2のみの学習・検証・テストデータを用いた場合の結果について示す.ま ず, Figs. 5.5-5.9 に学習における損失関数の推移を学習データ及び検証データのそれぞれに関 して , DNN の種類 (MLP , LSTM , GRU , BiLSTM , BiGRU) ごとにそれぞれ示す . いず れのモデルにおいても,100 エポックの学習で損失関数の十分な低下がみられ,検証データも 低下していることから過学習も回避できていることが分かる.次に , Figs. 5.10-5.14 にそれ ぞれ,スペクトログラム上に重ねる形でテストデータに対する予測値とラベルを,DNNの種 類(MLP , LSTM , GRU , BiLSTM , BiGRU) ごとに示す . これらの結果を見ると , いずれ のモデルも睡眠時のテストデータに対応する(f)から(i)は尤もらしい心拍が推定されている ことが確認できる. しかし, 日常活動時のテストデータである(a)から(e) はあまり良好な予 測ができなかった.特に, (b) , (d) ,及び (e) は 0 のラベルが付与されている予測困難なデー タであるが,予測結果は0にはならず低い心拍値で推移してしまっている.この問題は,予測 困難を別の質的ラベルではなく () という量的ラベルで付与してしまったことが原因と考えられ る. すなわち, 損失関数を平均二乗誤差で測っている以上, DNN は平均的な値を予測する傾 向になるので,0というラベルに対して高精度な予測が困難になる恐れがある.また,(c)及 び(j)にみられるような急なノイズに対する予測の精度も低いという結果になった.しかしな がら,この中では BiLSTM のモデルが比較的精度の良い予測を達成している.

次に,これらのテストデータの予測値に対して MAE 及び RMSE を算出した結果と,ピアソン相関を算出した結果をそれぞれ Fig. 5.15 及び Fig. 5.16 に示す.この結果より,MAE は BiLSTM が 0.23 Hz(14 bpm)程度と最も低い結果となった.しかしながら,この性能は心拍の推定としては十分な精度とは言い難い.また,Fig. 5.16 を見ると,ピアソン相関もいずれのモデルにおいても 0.7–0.8 程度となっており大きな差はないが,LSTM が最も高い結果となった.MAE に対してピアソン相関はそこまで低くなく,心拍の変動に対する追従性能はそ

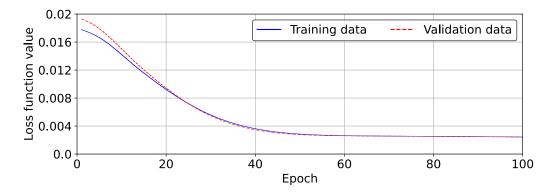


Fig. 5.5. Loss function values for training and validation datasets of Data 2 obtained by MLP.

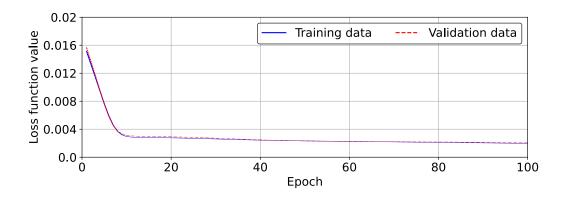


Fig. 5.6. Loss function values for training and validation datasets of Data 2 obtained by LSTM.

こまで悪くはないことが分かる.

5.4.2 全データを用いた学習と評価

本項では,データ1からデータ3のすべての学習・検証・テストデータを用いた場合の結果について示す.但し,ここではデータ1のテストデータに対する結果のみを掲載し,データ2及びデータ3のテストデータに対する結果については付録Cに掲載する.なお,2章で述べた通り,データ1からデータ3は測定条件が大きく異なるため,DNNの学習はその違いに対する汎化性能も獲得せねばならず,前項よりもより難しい学習となっている.

まず,Figs.~5.17-5.21 に学習における損失関数の推移を学習データ及び検証データのそれぞれに関して,DNN の種類(MLP,LSTM,GRU,BiLSTM,BiGRU)ごとにそれぞれ示す.前項のデータ 2 のみを用いた学習結果と同様に,いずれのモデルにおいても,100 エポックの学習で損失関数の十分な低下がみられ,過学習も回避できている.次に,Fig.~5.22-5.26 にそれぞれ,スペクトログラム上に重ねる形でデータ 1 のテストデータに対する予測値とラベルを,DNN の種類(MLP,LSTM,GRU,BiLSTM,BiGRU)ごとに示す.これらの結果

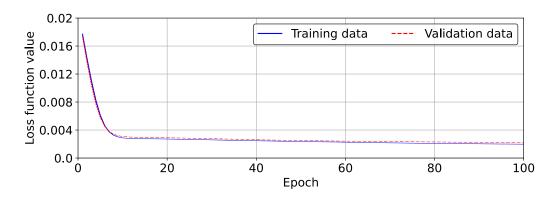


Fig. 5.7. Loss function values for training and validation datasets of Data 2 obtained by GRU.

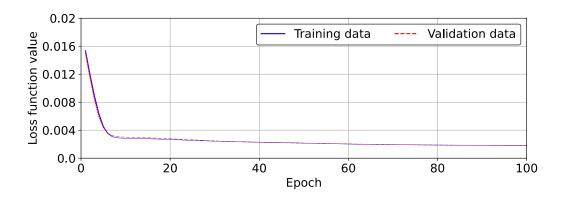


Fig. 5.8. Loss function values for training and validation datasets of Data 2 obtained by BiLSTM.

を見ると,データ 1 の日常活動時である (a) から (d) において,いずれの DNN も特に (a) と (d) である程度正確な予測ができていることが分かる.依然として (b) や (c) では推定に失敗してしまっているが,この結果はさらなる精度向上ができれば,日常活動時の PPG を用いて心拍の推定ができる可能性を示唆している.

次に,これらのテストデータの予測値に対して MAE 及び RMSE を算出した結果と,ピアソン相関を算出した結果をそれぞれ Figs. 5.27 及び 5.28 に示す.この結果では,データ 2 のみを用いた予測の結果よりも悪化しており,やはり PPG の測定条件に対する汎化性能を獲得することが非常に難しいという結論が導かれる.従って,次章で示すような改良により,提案手法の推定精度の全体的な改善が必要と考えられる.

5.5 本章のまとめ

本章では,DNNに基づく心拍推定法の実験を行い,テストデータに対する予測と評価について示した.5.2 節では,実験によって得られた心拍の推定値に適用する評価指標について説明した.5.3 節では,DNNの学習モデルの構造及び実験条件を示し,DNNで用いるデータの

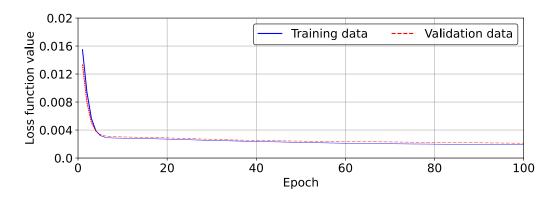


Fig. 5.9. Loss function values for training and validation datasets of Data 2 obtained by ${\rm BiGRU}.$

分割方法について説明した.5.4 節では、データ 2 のみを用いた実験及びすべてのデータを用いた実験のデータ 1 に対する心拍の予測結果及び評価について示した.次章では,本論文をまとめる.

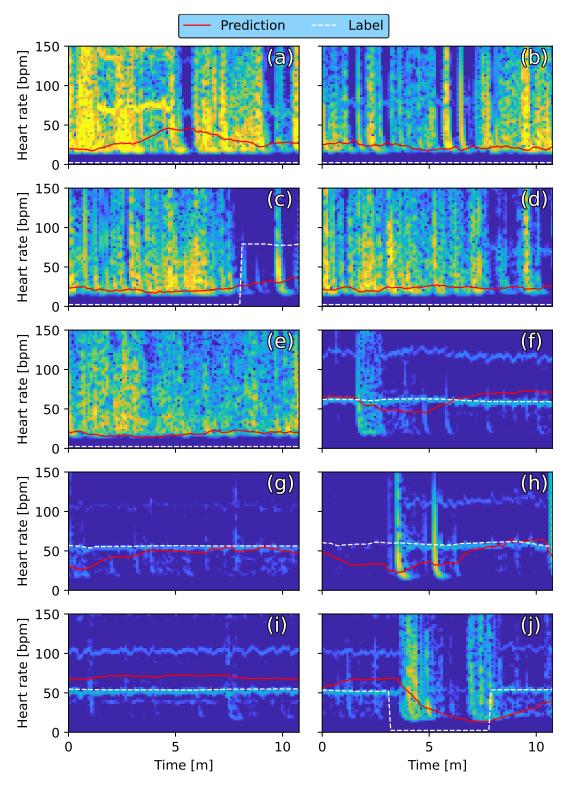


Fig. 5.10. Spectrograms and heart rates of labels and predicted results obtained by MLP trained with Data 2, where (a)–(j) correspond to each of test data.

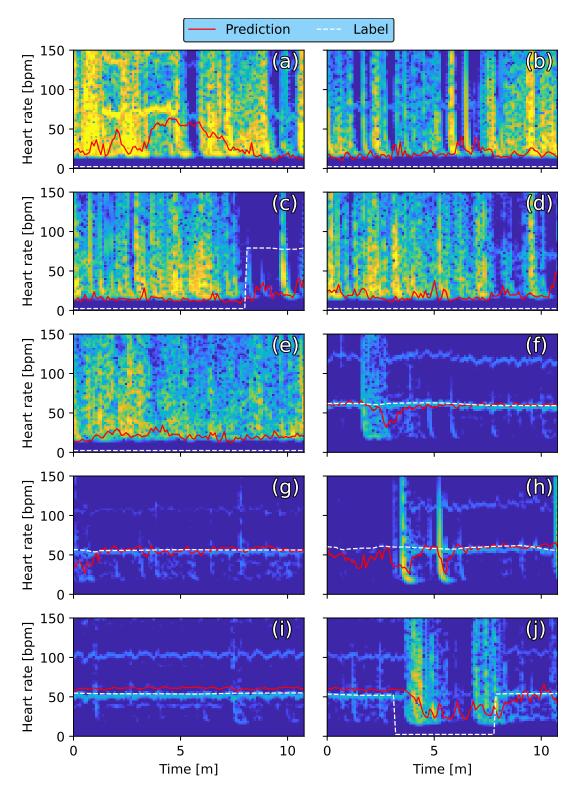


Fig. 5.11. Spectrograms and heart rates of labels and predicted results obtained by LSTM trained with Data 2, where (a)–(j) correspond to each of test data.

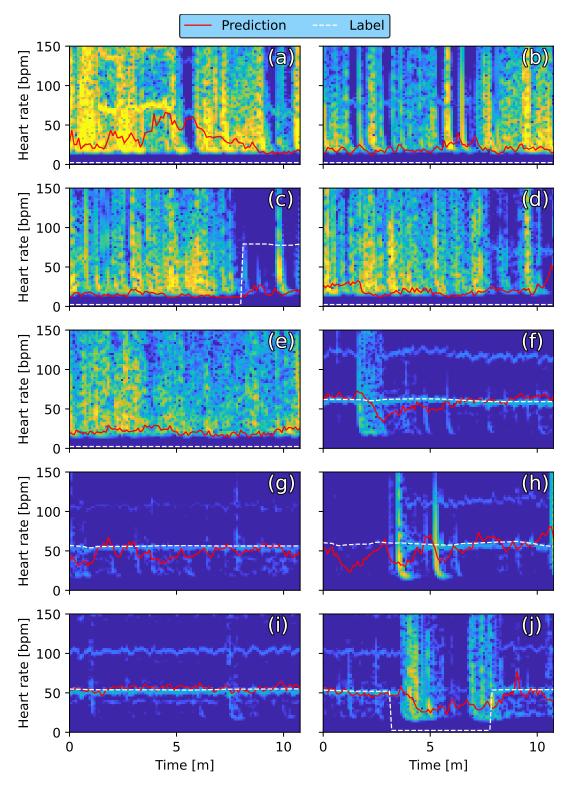


Fig. 5.12. Spectrograms and heart rates of labels and predicted results obtained by GRU trained with Data 2, where (a)–(j) correspond to each of test data.

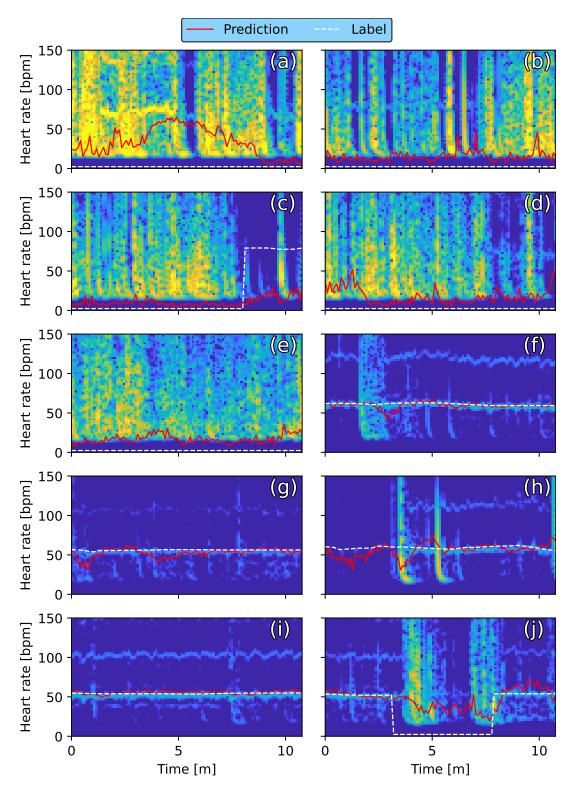


Fig. 5.13. Spectrograms and heart rates of labels and predicted results obtained by BiL-STM trained with Data 2, where (a)–(j) correspond to each of test data.

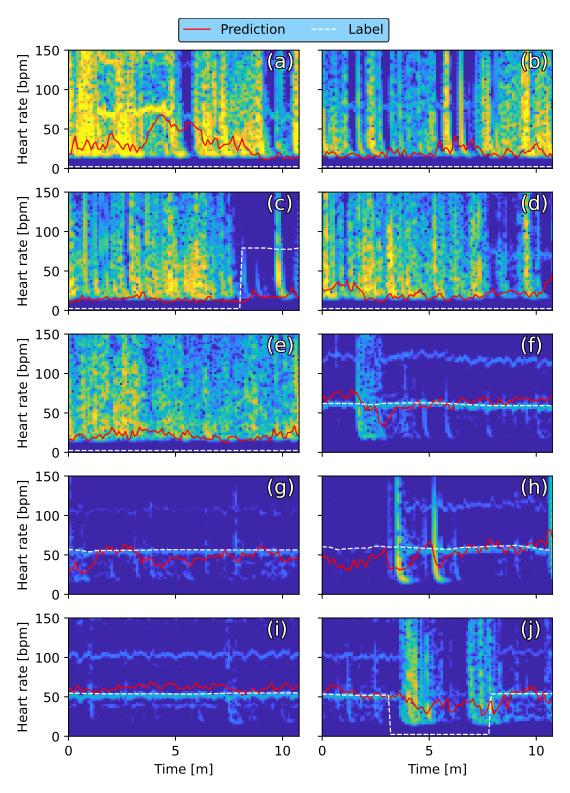


Fig. 5.14. Spectrograms and heart rates of labels and predicted results obtained by Bi-GRU trained with Data 2, where (a)–(j) correspond to each of test data.

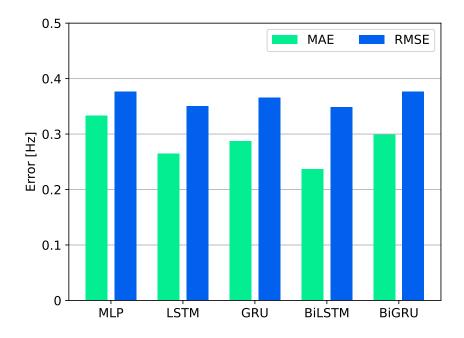


Fig. 5.15. MAE and RMSE of predicted heart rates for test data of Data 2 estimated by DNN models trained with Data 2.

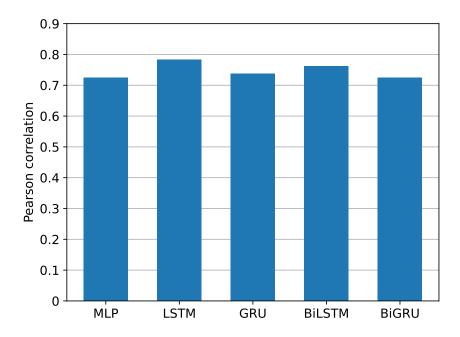


Fig. 5.16. Pearson correlation of predicted heart rates for test data of Data 2 estimated by DNN models trained with Data 2.

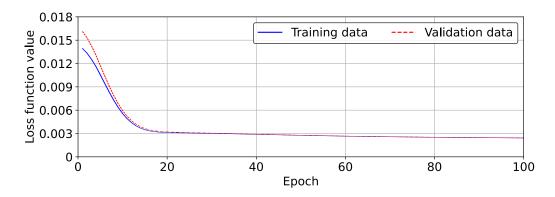


Fig. 5.17. Loss function values for training and validation datasets of all of Data 1, Data 2 and Data 3 obtained by MLP.

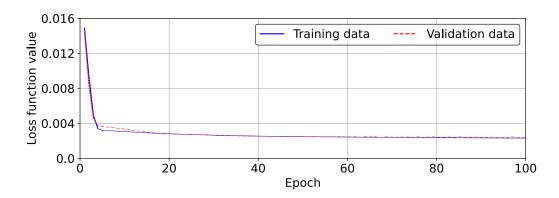


Fig. 5.18. Loss function values for training and validation datasets of all of Data 1, Data 2 and Data 3 obtained by LSTM.

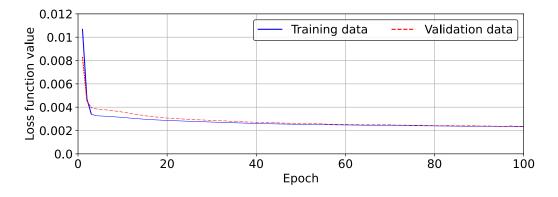


Fig. 5.19. Loss function values for training and validation datasets of all of Data 1, Data 2 and Data 3 obtained by GRU.

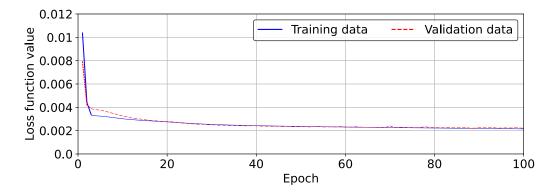


Fig. 5.20. Loss function values for training and validation datasets of all of Data 1, Data 2 and Data 3 obtained by BiLSTM.

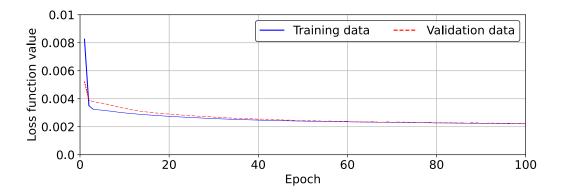


Fig. 5.21. Loss function values for training and validation datasets of all of Data 1, Data 2 and Data 3 obtained by BiGRU.

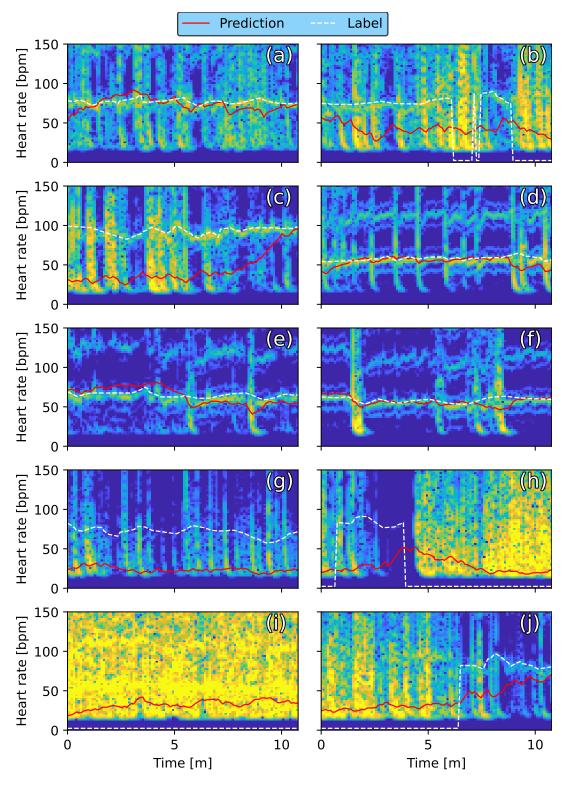


Fig. 5.22. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by MLP trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

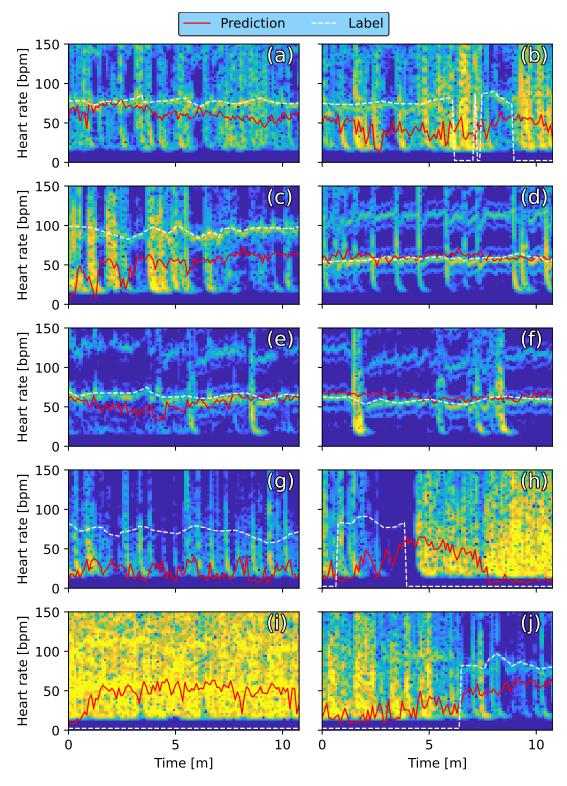


Fig. 5.23. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by LSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

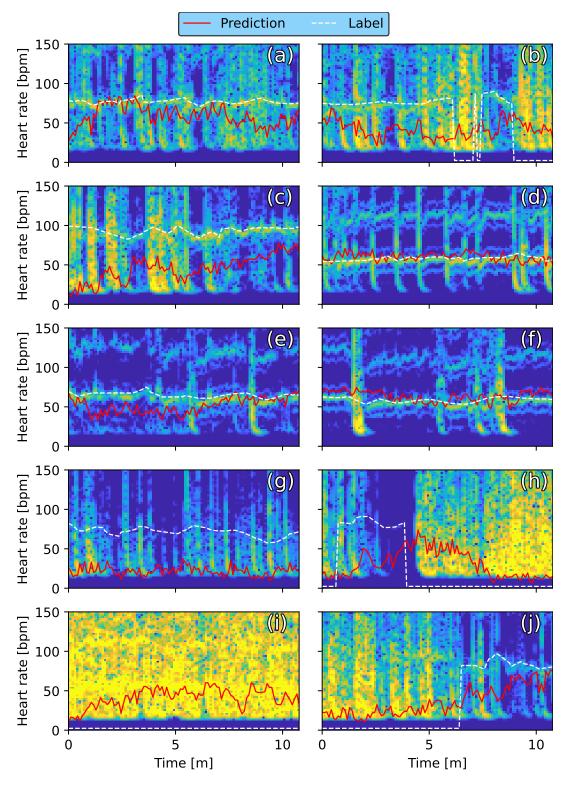


Fig. 5.24. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by GRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

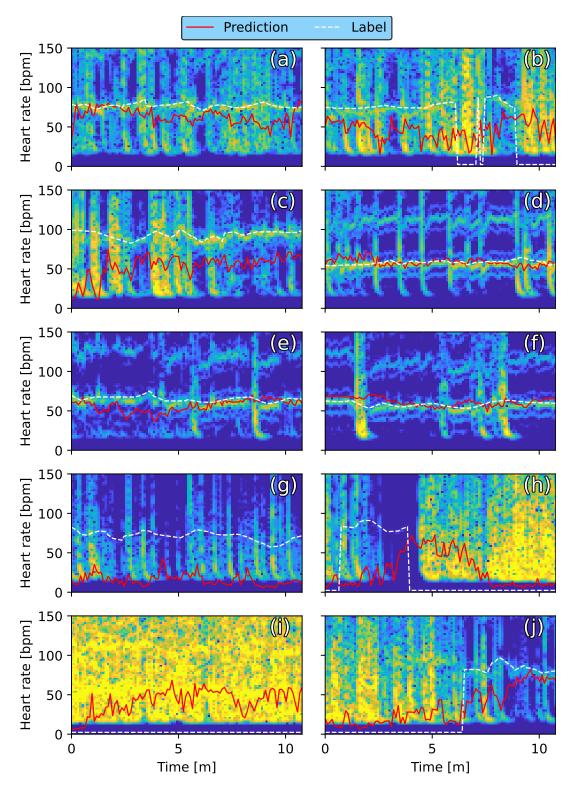


Fig. 5.25. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by BiLSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

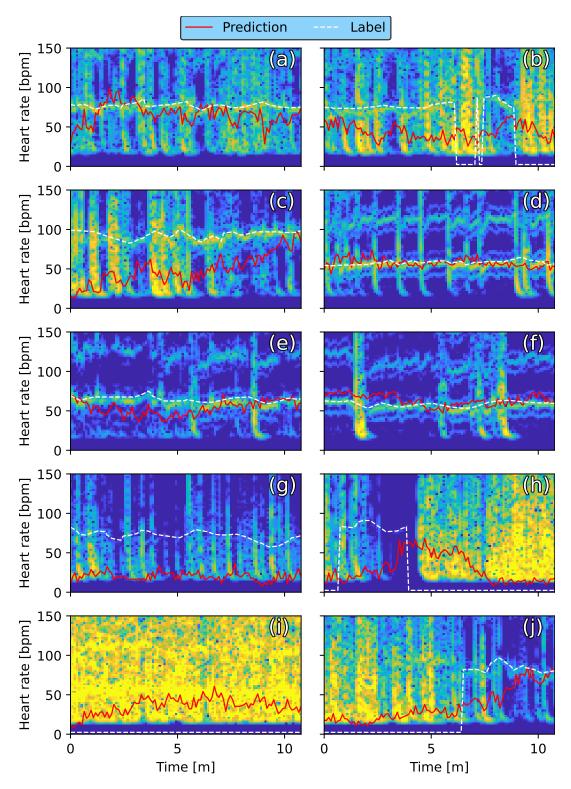


Fig. 5.26. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by BiGRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

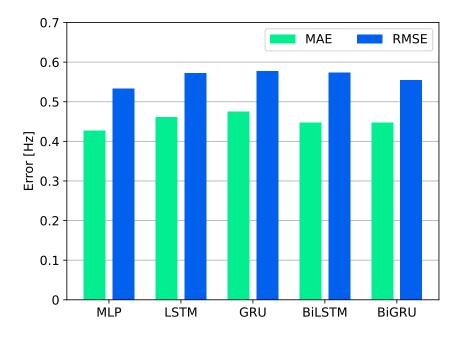


Fig. 5.27. MAE and RMSE of predicted heart rates for test data of Data 1 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.

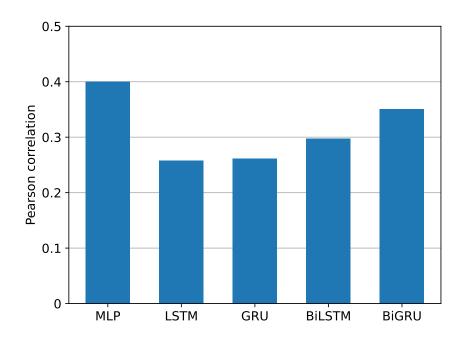


Fig. 5.28. Pearson correlation of predicted heart rates for test data of Data 1 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.

第6章

結言

本論文では,ニューラルネットワークに基づく付け爪型センサによって得られた PPG から 心拍を推定する手法を説明し実験を行った.アプリケーションを作成することで,DNN で用 いるうえで必要となる正解値を作成し,MLP,LSTM,GRU,BiLSTM,及び BiGRU の 5 つのモデルを用いて学習を行った.

実験結果より,1日分の測定波形のみを用いた実験では Figs. 5.15 及び 5.16 に示した結果より,5 つの DNN モデルの中で LSTM と BiLSTM が高精度な心拍推定結果を得た.3日分の測定波形すべてを用いた実験においても同様に,Figs. 5.27 及び 5.28 に示すように,MAE と RMSE は BiLSTM が最も良い値を示し,ピアソン相関は LSTM が最も良い値を示した. LSTM は入力,忘却,及び出力の 3 つのゲートで構成されており,時系列を考慮する学習において長期記憶を利用してモデルを構築する.今回の例では,PPG のスペクトログラムを DNN の入力としていて,これは時間方向に関して相関があるため LSTM と BiLSTM が有効であったと考える.しかしながら,3日分の測定波形すべてを用いた実験において,どの学習モデルにおいても精度が低下しているため,測定条件の統一が重要であるといえる.また,これらの結果より,DNN のモデルの種類については LSTM と BiLSTM の 2 つが他のネットワーク構造と比較して有効であった.

最後に今後の展望について述べる、教師あり学習は基本的に高精度なラベルが付与された学習データの量が増えるほど性能を改善できるため、さらに測定実験を繰り返して PPG を集めることが重要になる。また、今回は予測困難な時間フレームに対して 0 という量的な値を付与したが、これは 0 という値自身に特別な意味はないため不適切と考えられる.DNN のモデルを改良して、先に予測困難か予測可能かの分類問題を解く DNN を通し、予測可能な場合に心拍値の回帰を行う二段階のネットワーク構造を用いると、理論的には正しい予測が可能になる。そのような二段階のネットワークは、複数の損失関数を同時に最適化するマルチヘッドな DNN となるが、これを効率的に学習することができ、高い性能を達成できれば、本論文で扱った教師あり心拍推定のアプローチの妥当性がより明確になると思われる.

謝辞

本論文は,香川高等専門学校電気情報工学科北村研究室にて行われた研究に基づくものです. まず,本研究を進めるにあたり,ご多忙のところ熱心にご指導くださいました指導教員の北村大地講師に心より感謝申し上げます.北村大地講師には,論文執筆や研究に関する議論など,細部にわたるまで丁寧にご指導いただきました.

本論の副査である雛元洋一助教には、論文の構成や記述に関して大変有益な助言を頂き、大変お世話になりました、ここに厚く御礼申し上げます、

機械電子工学科の石井耕平准教授には、未経験の分野である生体信号処理の研究を行うにあたって、多数の助言と支援を賜りました、ここに感謝します、

北村研究室の先輩である専攻科 1 年の川口翔也氏・蓮池郁也氏・溝渕悠朔氏・村田佳斗氏には, MATLAB や Python といった不慣れな言語に関するアドバイス等をはじめ,数々のご支援をいただきました.また,北村研究室同期の唐渡昂希氏・岸本麗央氏・綾野翔馬氏には日頃のディスカッションをはじめ,研究室生活を様々な面で支えていただきました.ここに感謝申し上げます.

最後になりますが,現在に至るまで私の学生生活を金銭的に支え,暖かく見守って下さった 両親には感謝の念に堪えません.これまで本当にありがとうございました.

参考文献

- [1] Apple watch, https://support.apple.com/ja-jp/guide/watch/apda88aefe4c/watchos
- [2] ヘルスパッチ MD, https://www.omron.com/jp/ja/news/2015/11/h1105.html
- [3] K. Ishii and N. Hiraoka, "Nail tip sensor: toward reliable daylong monitoring of heart rate," *Trans. Electrical and Electronic Eng.*, vol. 15, no. 6, pp. 902–908, 2020.
- [4] L. Qiao, R. D. Mark, and G. D. Clifford, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter," *Physiological Measurement*, vol. 29, no. 1, pp. 15, 2007.
- [5] K. R. Arunkumar and M. Bhaskar, "Heart rate estimation from photoplethysmography signal for wearable health monitoring devices," *Biomedical Signal Processing and Control*, vol. 50, pp. 1–9, 2019.
- [6] D. Biswas, L. Everson, M. Liu, M.Panwar, B.-E. Verhoef, S. Patki, C. Kim, A. Acharyya, C. Van Hoof, M. Konijnenburg, and others, "CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment," *IEEE Trans. Biomedical Circuits and Systems*, vol. 13, no. 2, pp. 282–291, 2019.
- [7] A. Reiss, I. Indlekofer, P. Schmidt, and K. V. Laerhoven, "Deep PPG: Large-scale heart rate estimation with convolutional neural networks," *Sensors*, vol. 19, no. 14, pp. 3079, 2019.
- [8] L. Scalise and U. Morbiducci, "Non-contact cardiac monitoring from carotid artery using optical vibrocardiography," *Medical Engineering & Physics*, vol. 30, no. 4, pp. 490–497, 2008.
- [9] R. Špetlík, V. Franc, and J. Matas, "Visual heart rate estimation with convolutional neural network," in Proc. British Machine Vision Conference, pp. 3–6, 2018.
- [10] Z. Yu, X. Li, and G. Zhao, "Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks," arXiv preprint arXiv:1905.02419, 2019.
- [11] A. Mesleh, D. Skopin, S. Baglikov, and A. Quteishat, "Heart rate extraction from vowel speech signals," *Journal of Computer Science and Technology*, vol. 27, no. 6, pp. 1243–1251, 2012.

- [12] 梶谷奈美, 北村大地, 石井耕平, "付け爪型センサを用いたケプストラム分析及び MUSIC 法に基づく心拍推定," 電子情報通信学会 第 35 回信号処理シンポジウム, A1-4, pp. 5-10, 2020.
- [13] 梶谷奈美, "音楽信号処理における基本周波数推定を応用した心拍信号解析," 香川高等専門学校専攻科電気情報工学科コース 特別研究論文, 2022.
- [14] B. S. Atal, "Automatic speaker recognition based on pitch contours," J. Acoust. Soc. of Am., vol. 52, no. 6B, pp. 1687–1697, 1972.
- [15] IEEE Acoust., Speech, and Signal Process. Society, "Digital Signal Process. Committee," *Programs for Digital Signal Processing*, IEEE Press, New York, 1979.
- [16] A. Cheveigne and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," J. Acoust. Soc. of Am., vol. 111, no. 4, pp. 1917–1930, 2002.
- [17] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP–34, no. 3, pp. 276–280, 1986.
- [18] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in Proc. International Conference on Machine Learning, 2010.
- [19] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2016.
- [20] K. Cho, B. V. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv: 1406.1078, 2014.

付録 A

データ1のみを用いた学習と評価

本付録では、5章で掲載しなかったデータ1のみを用いた実験の結果を掲載する.Figs. A.1-A.7 にデータ1のみを用いた実験のテストデータに対する結果を示す.

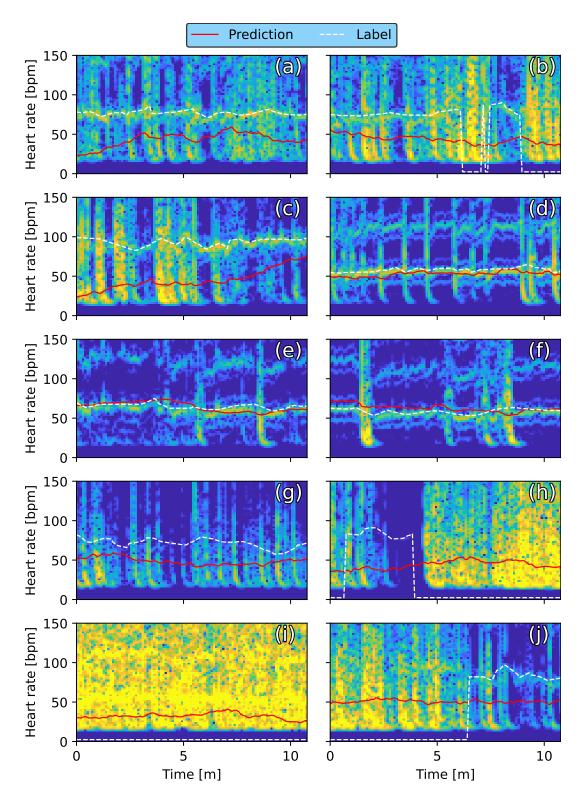


Fig. A.1. Spectrograms and heart rates of labels and predicted results obtained by MLP trained with Data 1, where (a)–(j) correspond to each of test data.

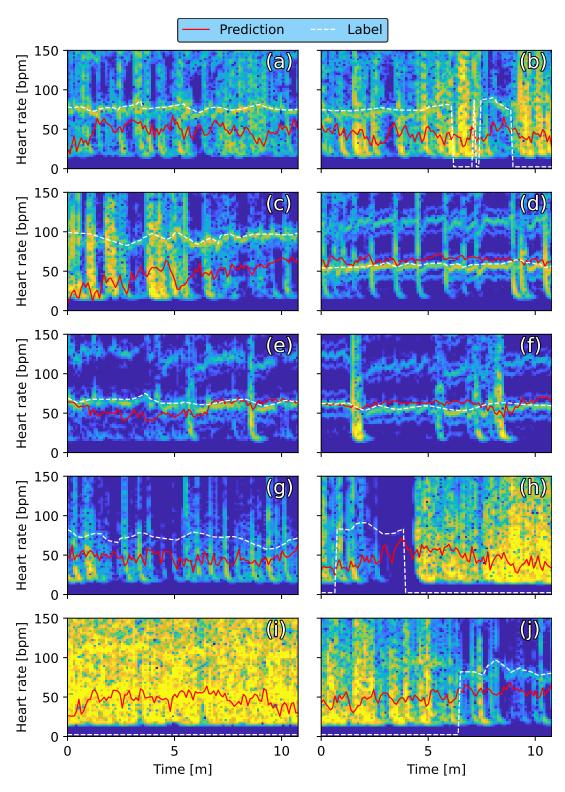


Fig. A.2. Spectrograms and heart rates of labels and predicted results obtained by LSTM trained with Data 1, where (a)–(j) correspond to each of test data.

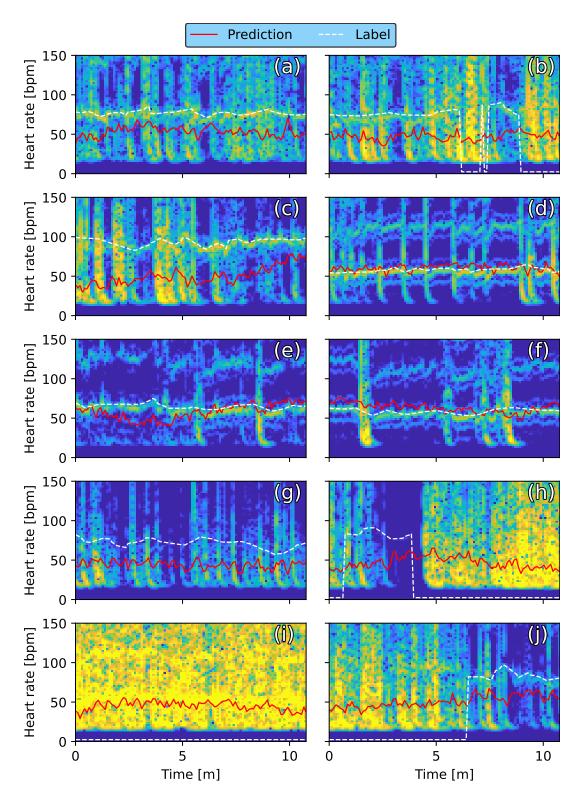


Fig. A.3. Spectrograms and heart rates of labels and predicted results obtained by GRU trained with Data 1, where (a)–(j) correspond to each of test data.

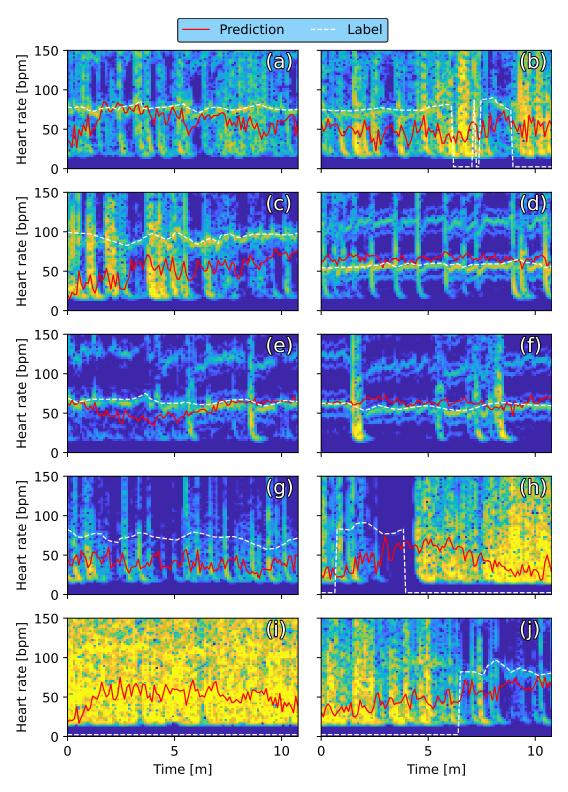


Fig. A.4. Spectrograms and heart rates of labels and predicted results obtained by BiL-STM trained with Data 1, where (a)–(j) correspond to each of test data.

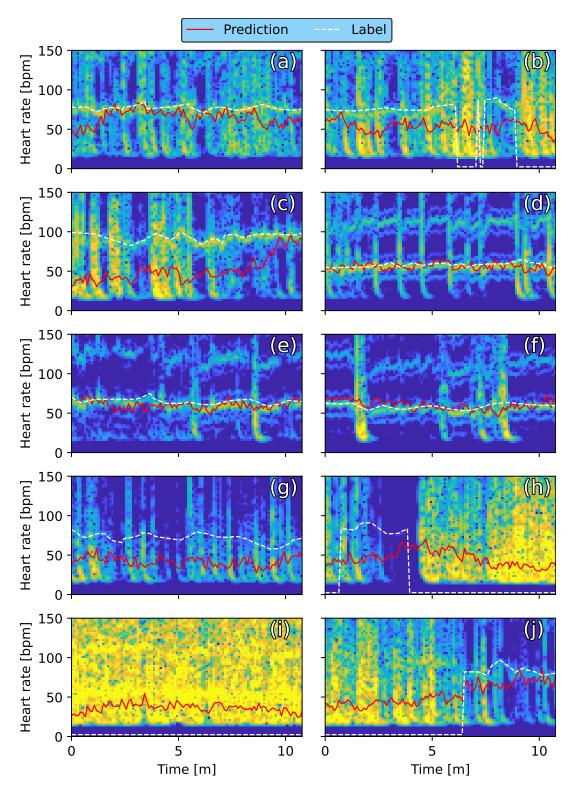


Fig. A.5. Spectrograms and heart rates of labels and predicted results obtained by BiGRU trained with Data 1, where (a)–(j) correspond to each of test data.

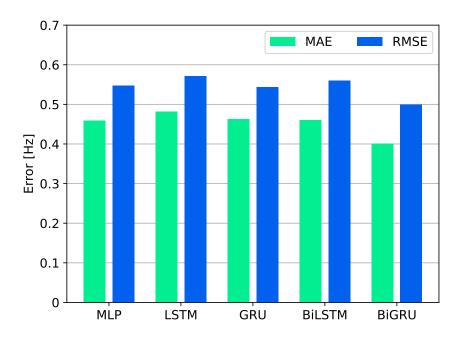


Fig. A.6. MAE and RMSE of predicted heart rates for test data of Data 1 estimated by DNN models trained with Data 1.

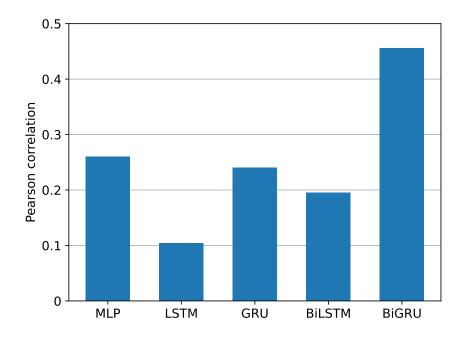


Fig. A.7. Pearson correlation of predicted heart rates for test data of Data 1 estimated by DNN models trained with Data 1.

付録 B

データ3のみを用いた学習と評価

本付録では、5章で掲載しなかったデータ3のみを用いた実験の結果を掲載する. $Figs.\ B.1-B.7$ にデータ3のみを用いた実験のテストデータに対する結果を示す.

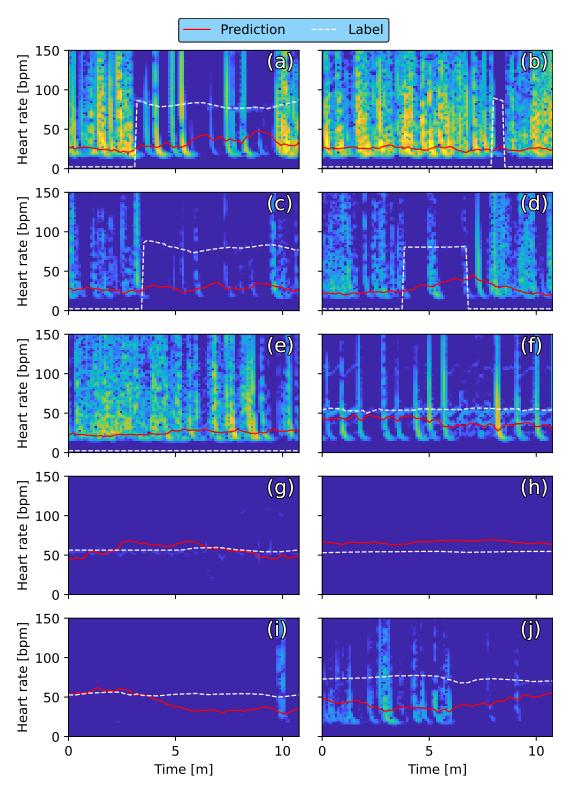


Fig. B.1. Spectrograms and heart rates of labels and predicted results obtained by MLP trained with Data 3, where (a)–(j) correspond to each of test data.

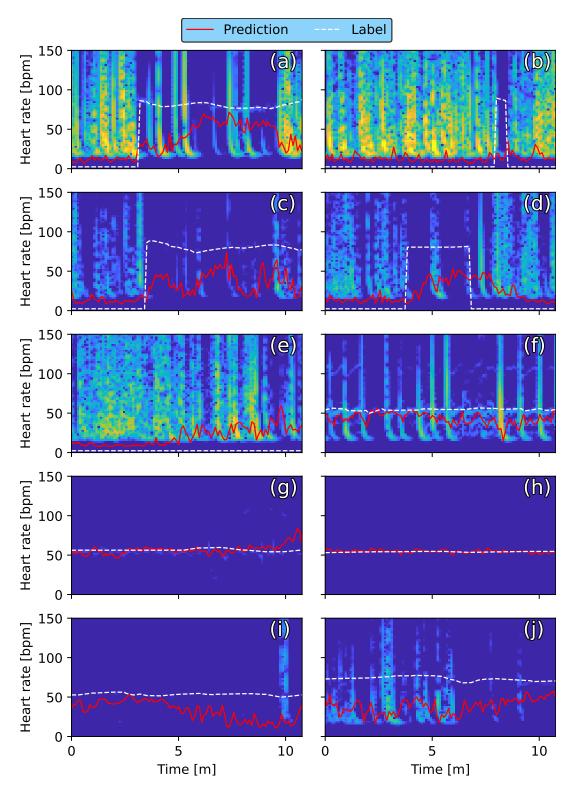


Fig. B.2. Spectrograms and heart rates of labels and predicted results obtained by LSTM trained with Data 3, where (a)–(j) correspond to each of test data.

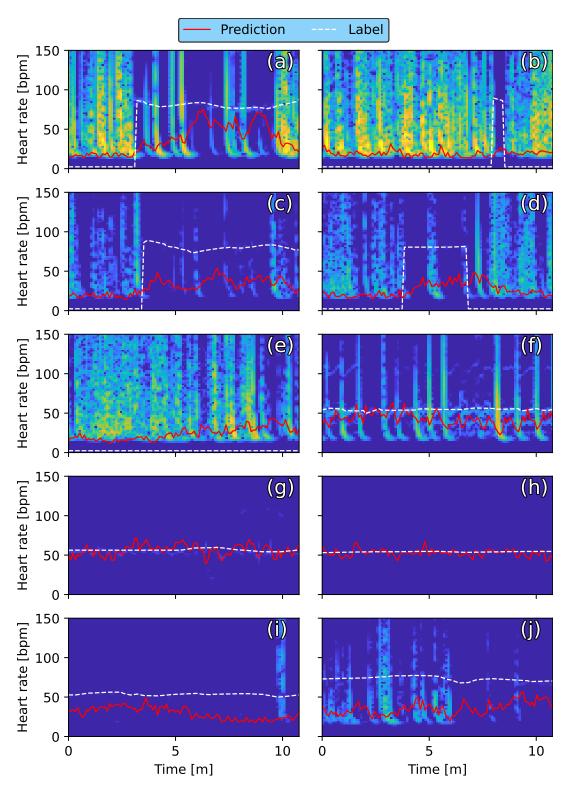


Fig. B.3. Spectrograms and heart rates of labels and predicted results obtained by GRU trained with Data 3, where (a)–(j) correspond to each of test data.

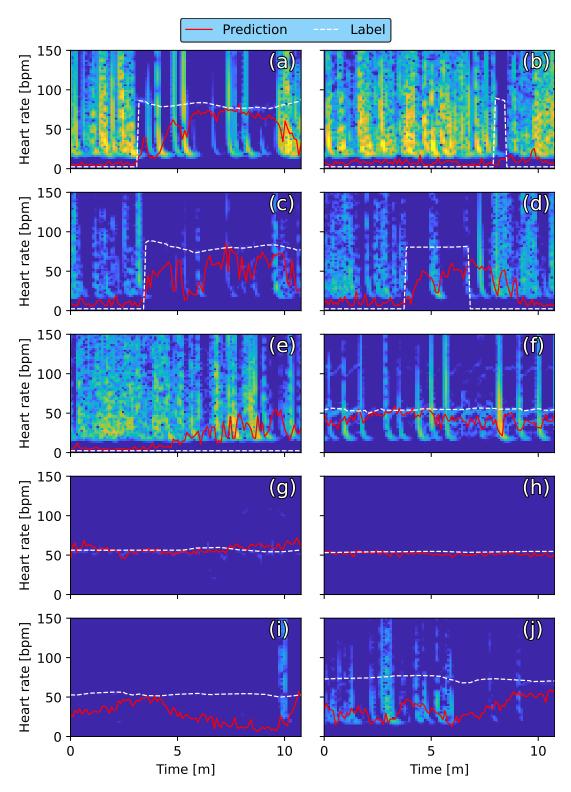


Fig. B.4. Spectrograms and heart rates of labels and predicted results obtained by BiL-STM trained with Data 3, where (a)–(j) correspond to each of test data.

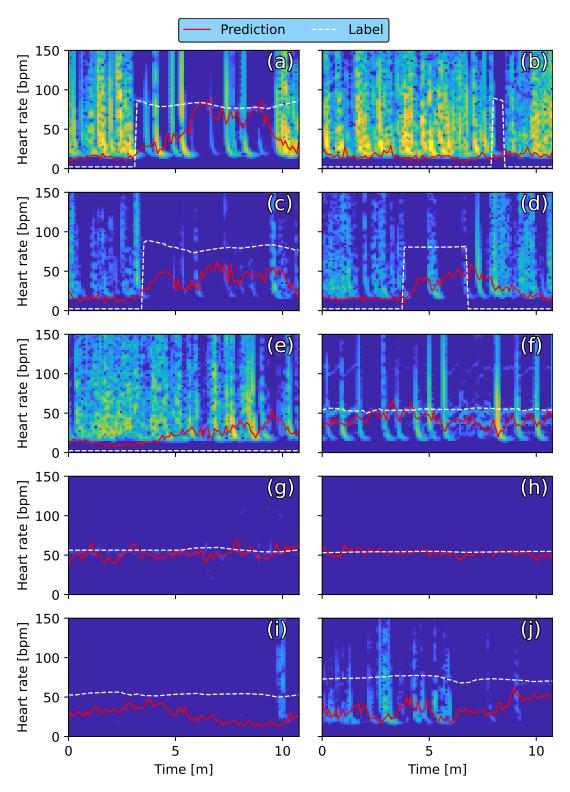


Fig. B.5. Spectrograms and heart rates of labels and predicted results obtained by BiGRU trained with Data 3, where (a)–(j) correspond to each of test data.

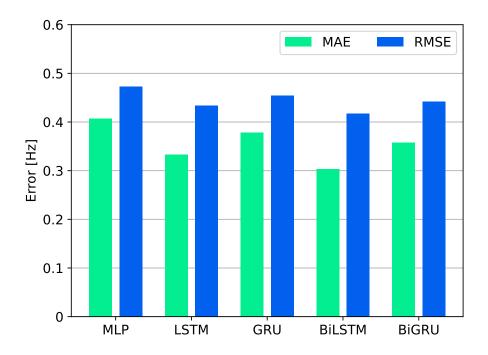


Fig. B.6. MAE and RMSE of predicted heart rates for test data of Data 3 estimated by DNN models trained with Data 3.

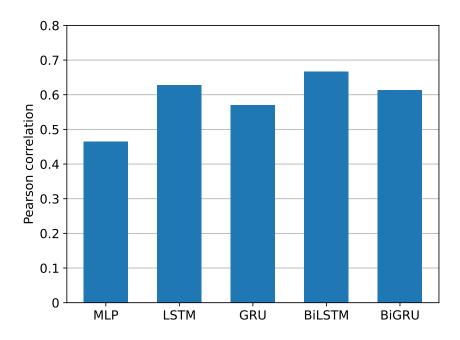


Fig. B.7. Pearson correlation of predicted heart rates for test data of Data 3 estimated by DNN models trained with Data 3.

付録C

全データを用いた学習とデータ2及 びデータ3に対する評価

本付録では,5章で掲載しなかったすべてのデータを用いた実験のデータ 2 とデータ 3 に対する結果を掲載する.Figs. C.1–C.7 にデータ 2 のテストデータに対する結果を示す.Figs. C.8–C.14 にデータ 3 のテストデータに対する結果を示す.

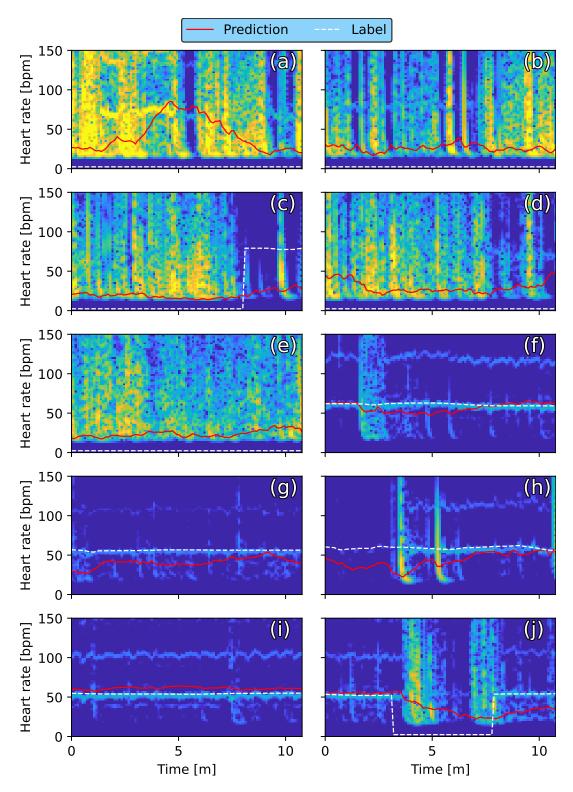


Fig. C.1. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by MLP trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

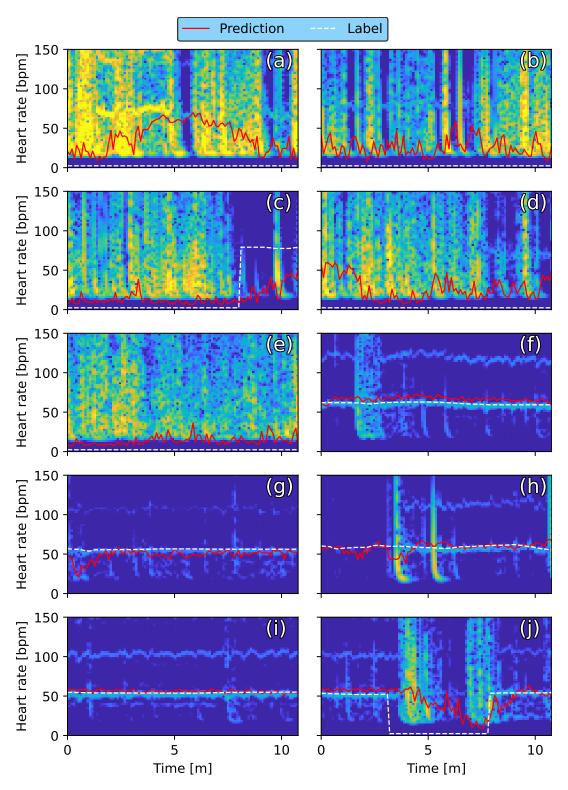


Fig. C.2. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by LSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

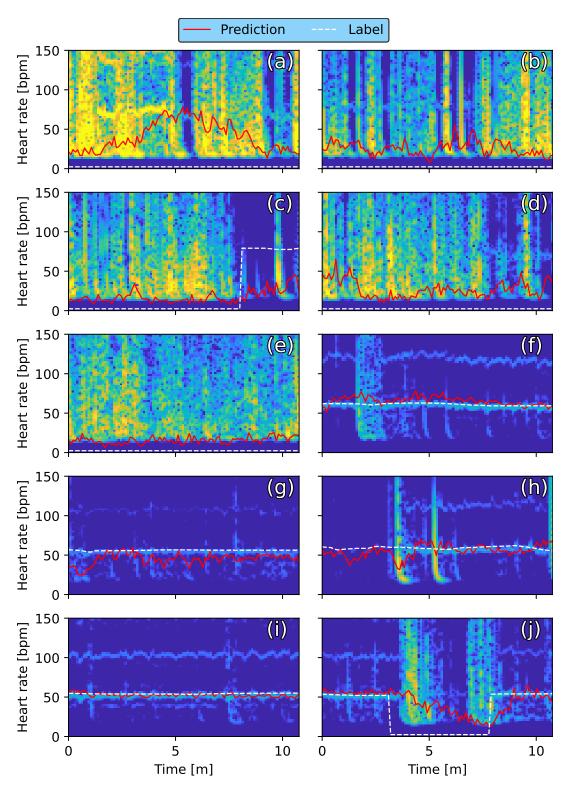


Fig. C.3. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by GRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

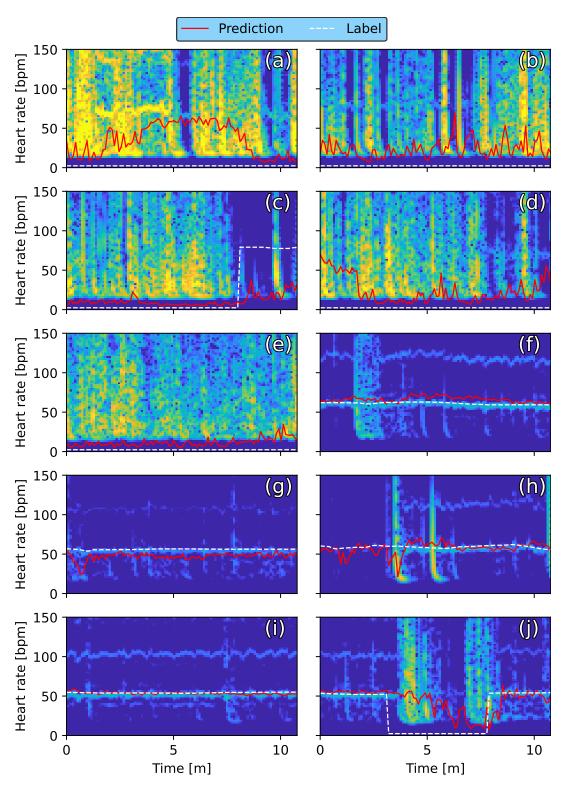


Fig. C.4. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by BiLSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

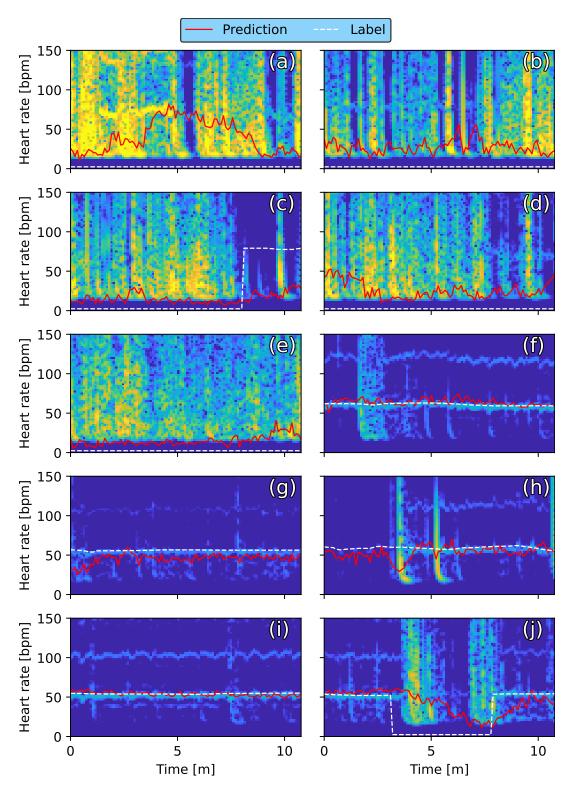


Fig. C.5. Spectrograms of Data 1 and heart rates of labels and predicted results obtained by BiGRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

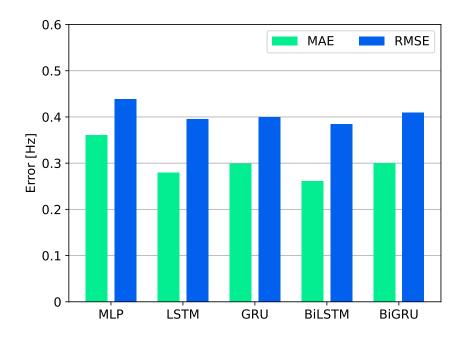


Fig. C.6. MAE and RMSE of predicted heart rates for test data of Data 3 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.

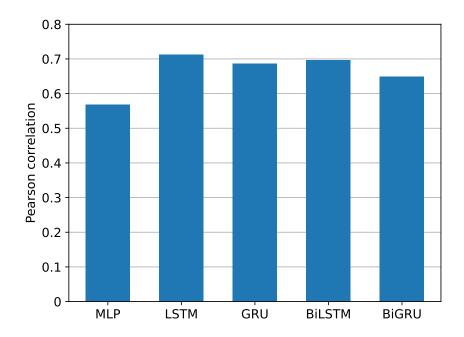


Fig. C.7. Pearson correlation of predicted heart rates for test data of Data 3 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.

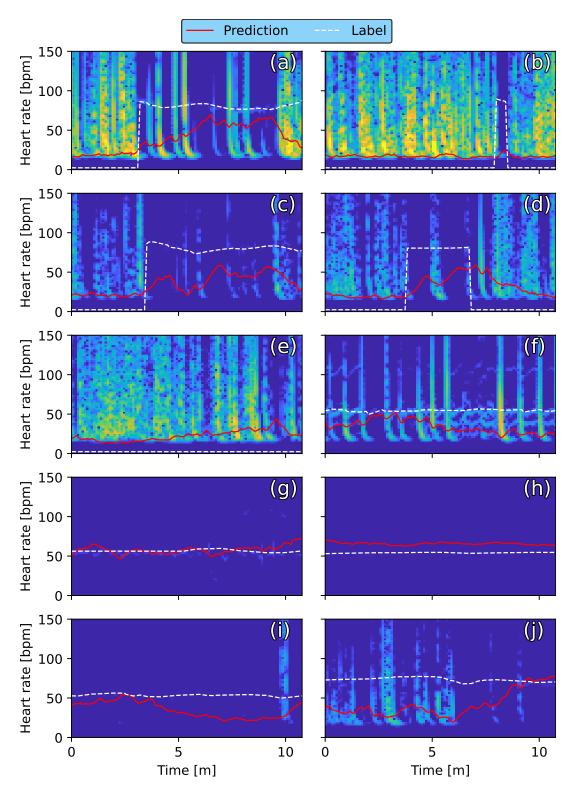


Fig. C.8. Spectrograms of Data 3 and heart rates of labels and predicted results obtained by MLP trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

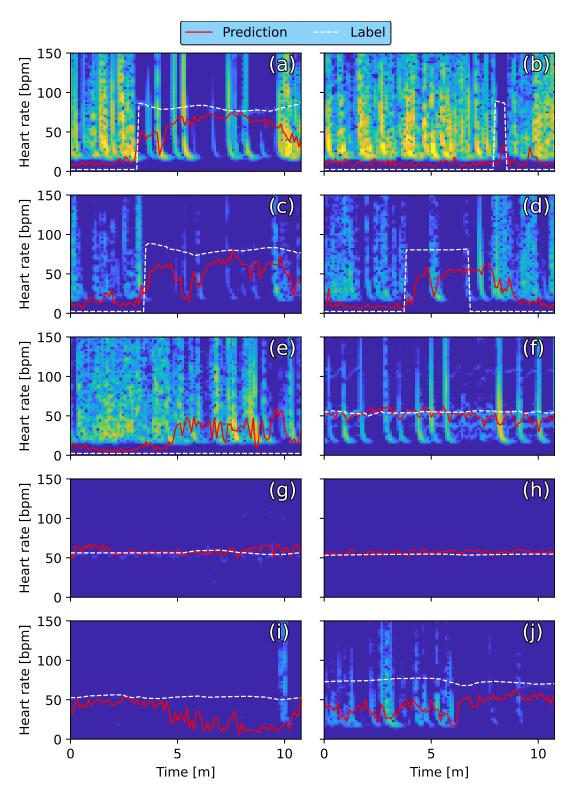


Fig. C.9. Spectrograms of Data 3 and heart rates of labels and predicted results obtained by LSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

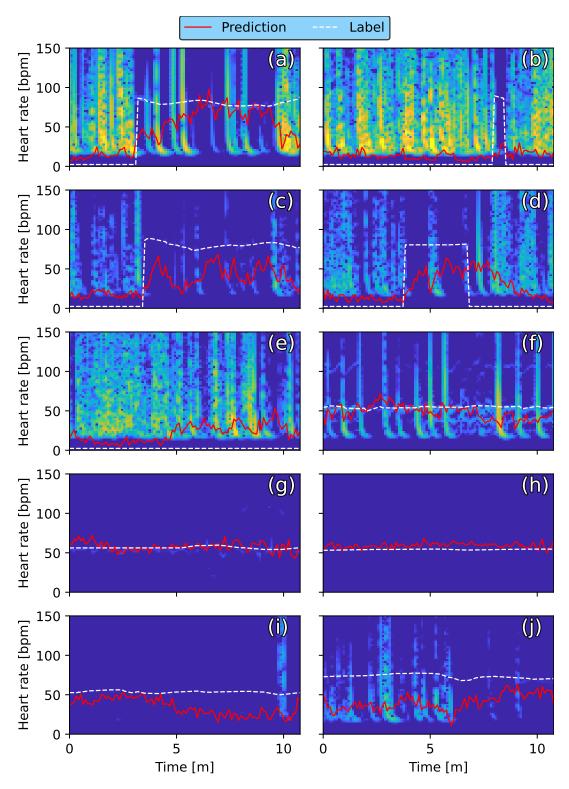


Fig. C.10. Spectrograms of Data 3 and heart rates of labels and predicted results obtained by GRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

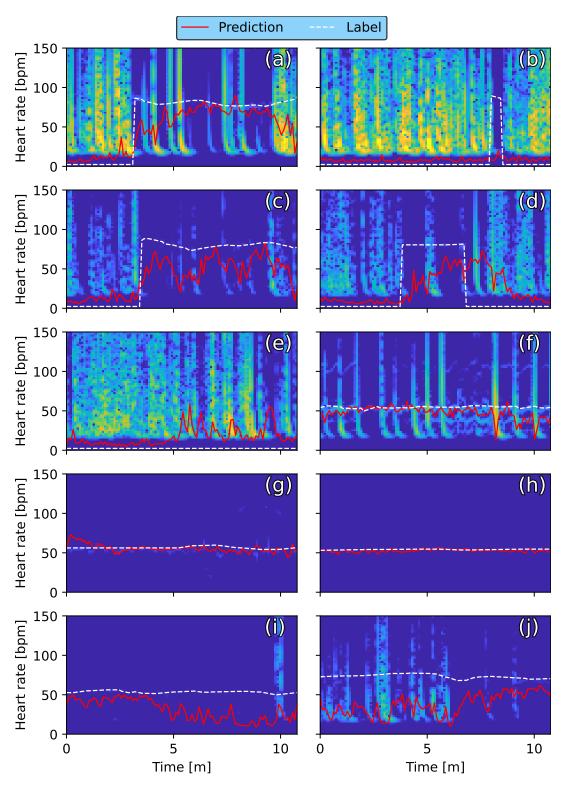


Fig. C.11. Spectrograms of Data 3 and heart rates of labels and predicted results obtained by BiLSTM trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

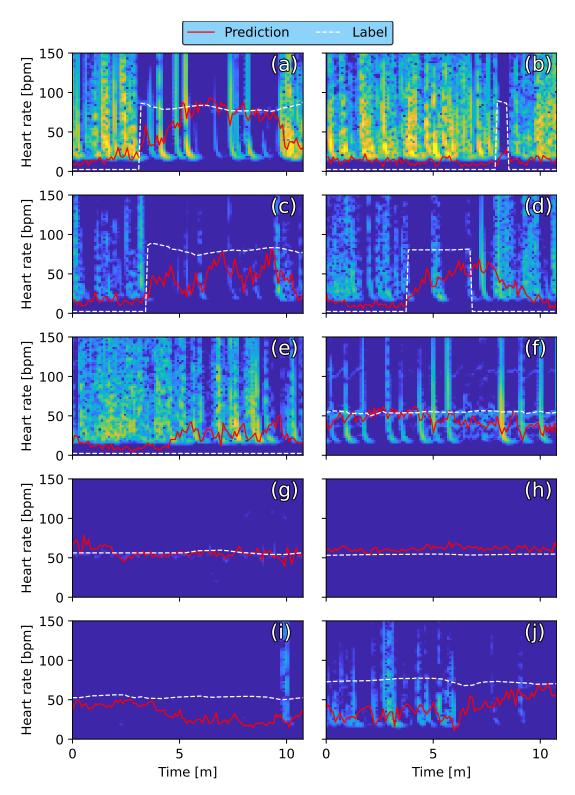


Fig. C.12. Spectrograms of Data 3 and heart rates of labels and predicted results obtained by BiGRU trained with all of Data 1, Data 2, and Data 3, where (a)–(j) correspond to each of test data.

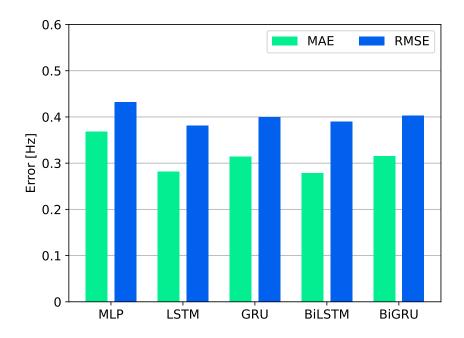


Fig. C.13. MAE and RMSE of predicted heart rates for test data of Data 3 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.

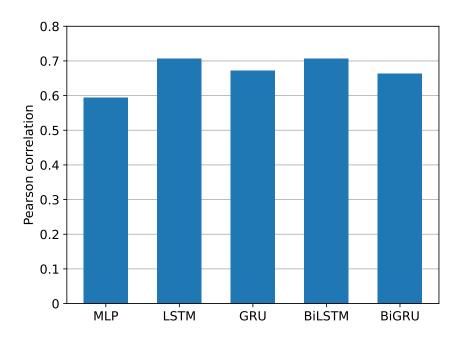


Fig. C.14. Pearson correlation of predicted heart rates for test data of Data 3 estimated by DNN models trained with all of Data 1, Data 2, and Data 3.