

畳み込みニューラルネットワーク及びLSTM を用いた異常心音の識別

市川嵩人¹ 桑原教彰¹

概要：近年、我が国では高齢化の進行に伴って、心疾患による死亡率が年々増加している。この傾向は今後も続くと予測されるため、心疾患に関する医療技術の発展が求められている。心疾患の初期診断方法である心音聴診は、医師の経験や勘などの個人に依存する要因に大きく左右される難しい診断である。一方、近年医療分野において、診断補助技術としての深層学習の活用が注目されている。本研究の先行研究では、深層学習を用いて正常心音と異常心音の機械的識別実験が行われ、平均95%以上の識別率が得られたことから、心音聴診の診断補助技術として深層学習がある程度有効である事が示唆された。しかしながら、先行研究の提案手法では、心音に突発的な異常が見られるタイプの症例に対して誤認識を起こすことが課題として残っていた。そこで本論文では、時系列データの学習に用いられる機械学習手法である、スライディングウインドウ法とLSTMを用いて、心音の機械的識別実験を行った。その結果、時系列を考慮することで心音の突発的な変化を識別できることが示唆された。

キーワード：心音聴診、心雜音、異常識別、深層学習、畳み込みニューラルネットワーク、LSTM.

A method of recognizing abnormal heart sound using CNN-LSTM network

SHUTO ICHIKAWA^{†1} NORIAKI KUWAHARA^{†1}

1. はじめに

1.1 日本社会の現状

近年、我が国の高齢化は世界でも類を見ないペースで進行している。内閣府が発表した我が国の人口動態統計統計[1]に基づいて、我が国と先進諸国の高齢化率の推移を比較すると、我が国は1980年代まで下位であったが、90年代から大幅に高齢化が進行し、2005年には世界で最も高い水準となっている。そしてこの傾向は今後も高水準を維持すると予測される。それに伴って、日本における死因も年々変化している。以下の図1は、昭和22年（1947年）から平成28年（2016年）の間における主な死因別死亡率の年次推移を表している[1]。死亡率が減少傾向にある多くの疾病に対して、心臓病（心疾患）による死亡率は年々増加傾向にある。昭和初期と現在とを比較すると、やはりがんの増加が顕著であるが、心疾患による死亡率も大幅に増加しており、現在では2番目に多い死因であることが分かる。

1.2 心疾患と診断方法

心疾患とは、心臓に関する疾病の総称である。例えば、よく耳にする不整脈は心疾患の一種であり、期外収縮や心室細動といった症状が存在する。現代医療における心疾患の診断は、基本的に聴診器を用いた心音聴診を行うことから始まる。心音聴診で何らかの異常が発見された場合、その病気や重症度を判断する為、心電図や心エコー検査、採血などの詳細な検査が行われる。初期診断の段階で異常が発見されなければ、その後の詳細な検査も行われないため、心音聴診で異常を検知することは非常に重要である。しかしながら、心音の微細な異常を人間の聴覚のみで聞き分ける必要のある心音聴診は、熟練の経験や勘を必要とし、難しい診断である。

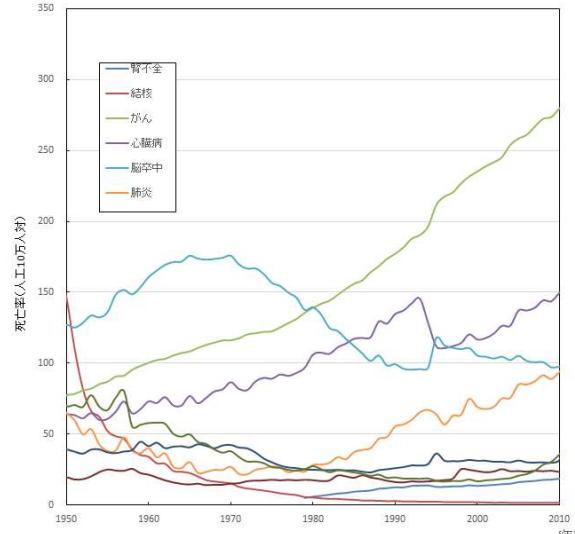


図1 日本における死因別死亡率[1]
Figure 1 Cause-specific Mortality in Japan

1.3 心音

心音とは、心臓が鼓動する際の活動に伴う、弁の開閉や血液の流れによって生じる音である。心臓の1回の心拍は、心臓が収縮して体全体に血液を送り出す収縮期と、心臓が拡張して心臓内に血液を取り込む拡張期に分けられる。心音は基本的にI音、II音と呼ばれる二つの音からなっており、I音は収縮期の開始時に、II音は収縮期から拡張期への移行時に発せられる。心臓に何らかの疾患があった場合、III音やIV音、心雜音と呼ばれる異常な心音が聴取されるようになる。前述の通り、基本的に心音はI音とII音から成るが、個人によっては何らかの原因で心音と同時、あるいは

心音の間に雑音が混じる場合がある。この雑音を、医学では経験的に心雜音と呼んでおり、聴取される心雜音の特徴から病気の推定を行う。心雜音は発生するタイミングにより、収縮期雑音、拡張期雑音、連続性雑音の3種に大別される。拡張期雑音、連続性雑音はすべて病的であるが、収縮期雑音には病的なものと病的でないものがある。病的でない心雜音は機能性心雜音と呼ばれる。実際の診断の際に心雜音の有無だけではなく、病的か否かを正確に診断する必要がある。このように、心雜音は心音の正常と異常を判別する上で重要な情報を含んでいると言える。

1.4 深層学習

深層学習とは、人間の脳を模した学習アルゴリズムである、ニューラルネットワークを多階層に重ねた機械学習手法である。深層学習には、入力データの特微量をコンピュータ自らが獲得するという特徴がある。本研究ではこの特徴に着目し、深層学習を用いることで、心音聴診において重要な「医師の長年の経験や勘」に依存することなく、心音データから必要な特微量を獲得し異常の識別が出来るのではないかと考えた。本研究では、畳み込みニューラルネットワーク（Convolutional Neural Network：以下、CNN）とLSTM（Long Short Term Memory）の2種類の深層学習モデルを使用した。

1.5 畳み込みニューラルネットワーク

CNNとは、何層もの深い層を持つニューラルネットワークで、画像認識の分野で優れた性能を発揮する学習モデルである。CNNは「畳み込み層」と「プーリング層」という2つの特徴的な層があり、畳み込み層は、入力画像のある領域内の値を合計して、特徴マップを作成する。これはエッジ抽出など、特徴抽出の役割を果たしている。プーリング層は、データの次元削減を行い、平行移動や回転など、特徴マップの位置ずれに対して頑強性を与える役割を持っている。こうして、CNNは画像の特徴を読み取り、位置の変化に惑わされることなく、画像の識別を行うことが出来る[2]。

1.6 LSTM

LSTM（Long Short Term Memory）は系列データの学習に用いられる機械学習の学習モデルである。最大の特徴として、従来のRNNでは認識不可能であった、「長期依存」を認識可能であることが挙げられる。文章や対話の生成、株価の予想などに利用される。本論文では、心音の時間的変化を認識することで、異常を識別出来るのではないかと仮説を立てた。

1.7 関連研究

医療分野において、深層学習が応用される例は増えている。深層学習は画像認識に優れた技術であるため、X線写真や核磁気共鳴画像（MRI）などの医療画像に対して深層学習を応用した研究例が多い。例えば、脳のMRIスキャン画像とその脳年齢をコンピュータに学習させ、未知のMRIスキャン画像の脳年齢を3秒で判定する研究や、胸部X線写真から肺の異常を検知し、医師の診断補助を行う研究等がある。これらの研究例は双方とも高い判定精度を得ており、医用画像の利用可能性が高いことを示唆している。

1.8 先行研究

本研究の先行研究では、北田らによって心音の機械的識別実験が行われた。北田らは、心疾患の初期診断として重要な心音聴診に着目し、医療分野でも注目されている機械学習の手法を用いて、診断補助技術としての深層学習の有効性を検証した。

先行研究における識別実験では、CNNを用いて、心音データの学習と未学習データを用いたテストによって学習モデルの評価を行った。心音データに機械学習を適用した

関連研究において、音声を画像データに変換して深層学習と組み合わせることで、従来の機械学習手法に比べて心音の識別精度が向上したと報告されている[3]ため、心音を3次元画像データに変換した。実験の結果、正常心音と異常心音の識別において平均95%以上の識別率を記録し、医師の診断補助技術として深層学習がある程度有効であることが示唆された。しかしながら、実験毎に誤認識を起こした心音データを分析すると、「心音に突発的な異常が見られる」パターンの症例に対して、北田らの提案手法は識別率が低いことが判明し、これを課題として今後改善することを検討していた[4]。

1.9 研究目的

本研究の目的は、心疾患の診断補助技術として、機械学習による心音の異常識別をより正確に行うことである。先行研究において北田らは、深層学習を心音データに適応した研究例が少ないと着目し、異常心音の機械的識別実験を行った。その結果、心音の正常異常識別において一定の識別率を記録し、心音聴診の診断補助技術として深層学習が有効であることを明らかにした。しかしながら、心音に突発的な異常が起る、期外収縮や心拍停止などの症状に対して、北田らの提案手法は誤認識を起こした。そこで本論文では、その課題に対して、心音を時系列データとして人工知能に認識させることで、突発的に発生する異常心音のパターンを識別できるのではないかと仮説を立て、機械学習において時系列データの認識が可能である学習モデルを用いて、心音の識別実験を行った。

2. スライディングウインドウ法を用いた識別実験

2.1 実験概要

本実験では、スライディングウインドウ法に基づいて製作したデータセットを用いて、深層学習による心音の異常識別を行った。本実験の目的は、期外収縮や心拍停止などの、心音に突発的な異常が見られる症例をコンピュータが正しく認識し、総合的に高い識別率を得ることである。また、本実験には学習モデルにCNNのみを採用しており、4章で行うCNN-LSTMによる識別実験との比較対象としての役割もある。実験に使用する心音データは医学書に付属するサンプル心音と、健康な実験協力者から採取した採取心音を用いた。これらの心音データは画像に変換し、スライディングウインドウ法を用いて時系列にトリミングした後、正常と異常のラベル付けを行ってデータセットとした。データセットは訓練用とテスト用に分割し、コンピュータによる識別テストを行ってモデルの評価を行った。

2.2 実験概要

本実験の大まかな手順を以下に示す。

1. 心音データをスペクトログラム画像に変換する。
2. スペクトログラム画像をスライディングウインドウ法に基づいてトリミングする。
3. 画像データを正常な心音と異常な心音に分け、それぞれ別のフォルダーに格納することでラベル付けを行つてデータセットを10セット作成する。
4. コンピュータによる学習とテストを10セット行う。
5. 識別率と損失値から学習モデルの評価を行う。

2.3 心音データ

本実験で使用する心音データは、医学書に付属するサンプル心音と、先行研究において実験協力者から採取された、採取心音がある。前者に関しては、『電子聴診器でぐんと身につく心音聴診技術』[5]、『聴診の達人』[6]に付属する心音データを引用した。後者は先行研究において、北田らが、聴診器とボイスレコーダーを組み合わせた自作の

実験装置を用いて採取したものである。聴診器は、FOCAL社製FC201の先端部分を使用し、ボイスレコーダーはオリンパス社製LS-7を使用している。心音の採取に際しては、医師の方の意見のもと、ベル面を心尖部に押し当てて採取された。なお本論文で用いる心音の聴診部位は、すべて心尖部のみを対象としている。これは、本研究において、異常心音を識別する上で重要な情報となるIII音が心尖部のみで聴取される音であり、人工知能による異常識別の有用性を考える上で最も適切な部位であると考えられたからである[4]。これに合わせるために、本論文では、医学書から採取した他の心音データもすべて心尖部のものを用いている。採取実験の被験者は、いずれも健康な20代の男性である。採取された心音データは、被験者の体動や、マイクと聴診器の接合部から生じる外部ノイズが含まれるため、このノイズを除去する音声処理が行われている。音声処理ソフトウェアであるAudacityを用いて、外部ノイズのみの音声ファイルをフィルターとすることで、外部ノイズが除去されている。また、他の心音データとのレベルの差異を無くすことと、心音をより明瞭にするために、ノーマライズ処理も行われている。なお、サンブル心音に関しては、外部ノイズと思われるノイズが比較的少ないため、ノイズ除去処理は行っていない。

2.4 画像データ変換

心音データを深層学習させるに際して、スペクトログラム図と呼ばれる画像データに変換する。ここで言うスペクトログラム図とは、音声信号に窓関数をずらしながら掛け合わせ、高速フーリエ変換することで、音の周波数特性を表現したものである。一般に、音声信号の周波数や位相の時間変化などを解析するために用いられる。以下の図2に、作成した正常心音のスペクトログラム図の例を示す。



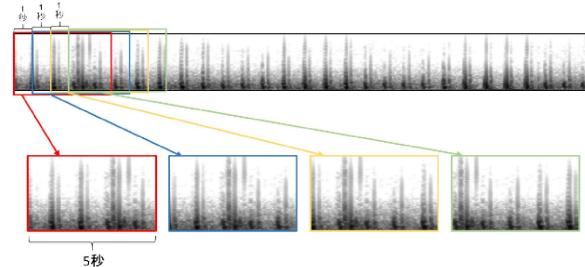
図2 正常心音のスペクトログラム図

Figure 2 Spectrogram Image Data of Normal Heart Sound

この画像は学習に使用するために軸や軸ラベルを除いていいが、縦軸が周波数、横軸が時間、色の濃さが音の強さを示している。また、データの容量を小さくし、学習の処理速度を向上させるため、グレースケールでプロットしている。

2.5 スライディングウィンドウ法

スライディングウィンドウ法とは、ある特定の時間的に連続した入力データに対して窓関数を当てはめ、入力データを部分系列に分割する手法である。行動認識の分野では、身体に装着した加速度センサなどから得られた時系列データに対して適応し、コンピュータに「行動の変化点」を認識させるために用いられる[7]。本研究では、このスライディングウィンドウ法を用いることで、不整脈や心拍停止などの「突発的な変化点」が存在する心音データを正しく認識することが可能になると仮説を立て、識別実験を行った。本実験では、窓関数の幅を5秒、オーバーラップ（ずらし幅）を1秒としてスライディングウィンドウ法を心音画像データに適応し、1秒ずつずれた長さ5秒の心音画像データを大量に作成した。以下の図3にスライディングウィンドウ法の模式図を示す。

図3 スライディングウィンドウ法
Figure 3 Sliding Window Method

2.6 データセット

本実験では、用意した心音の画像データを正常心音と異常心音の2種類に分け、コンピュータによる識別実験を行った。この際に用いた正常心音の中には、正常心音と正常心雜音の両方が含まれている。

また、すべての識別実験に関して、用意した心音画像データのうち、約10%をランダムに選択してテスト用データとし、残りの約90%を訓練用データとして用いた。なお、テスト用データと訓練用データの両方にすべての症例が含まれるようにするために、各症例に関して、それぞれ10%をランダムに選択してテストデータとしている。また、訓練用データのうち約10%をバリデーション用いた。本実験では10回の識別実験を行うため、このデータセットを10セット用意した。データセットの画像の枚数を以下の表1に示す。

表1 データセットの詳細
Table 1 Dataset Details

	正常心音	異常心音	計
学習用	1315	870	2185
テスト用	150	115	265
計	1465	985	2450

2.7 学習モデル

学習モデルには畠み込みニューラルネットワークを用いた。実装にはpythonのディープラーニングのフレームワークであるKerasを使用した。以下の表2に学習モデルの詳細を示す。オプティマイザ（最適化アルゴリズム）は、先行研究で効果を発揮したAdamを採用し、学習率等のパラメータはAdamを提案した論文[8]に従って採用された、デフォルトの値を用いた。

表2 CNNモデルの層構成
Table 2 CNN Model

Layer	Kernel	Outputs	Function
input		$3 \times 25 \times 25$	
conv1	3×3	$3 \times 25 \times 32$	relu
conv2	3×3	$3 \times 25 \times 32$	relu
pool1	2×2	$2 \times 13 \times 32$	max
conv3	3×3	$2 \times 13 \times 64$	relu
conv4	3×3	$2 \times 13 \times 64$	relu
pool2	2×2	$1 \times 7 \times 64$	max
conv5	3×3	$1 \times 7 \times 128$	relu
conv6	3×3	$1 \times 7 \times 128$	relu
pool3	2×2	$1 \times 4 \times 128$	max
flatten		512	
dense1		512	relu
dense2		2	softmax

2.8 実験結果

以下の図表はコンピュータによる学習の経緯と結果を示したものであり、図4は識別率、図5は損失値を表したグラフである。10セットの実験において、すべて同様のグラフが出力された。各図において、青線は、訓練用データによって学習を行い、訓練用データによってテストを行った際の識別率である。また、オレンジ色の線はバリデーション（検証）用データでテストを行った際の識別率である。最終的な評価はテスト用データを用いたテストの識別率で行う。表3に識別率と損失値及びその平均を示す。また、表4に誤認識を起こした症例のリストを示す。

表5の通り、本実験においては、結果として平均98%以上の高い識別率を得ることが出来た。図4から、学習回数を重ねる毎に識別率が向上し、高い識別率で頭打ちとなっていることが見て取れる。基本的に訓練データによるテストの識別率（青線）はほぼ100%の値を推移しており、バリデーションデータによるテストの識別率（オレンジ線）は、0~75エポックの間では振動しながらも、やがて振れ幅が小さくなり、高い水準に落ち着いている。また、図5を見ると、学習回数を重ねるにつれて、損失値が減少する傾向が見て取れる。これらから、本実験では適切に学習が進んだことにより、良好な識別率を得ることが出来たと考えられる。

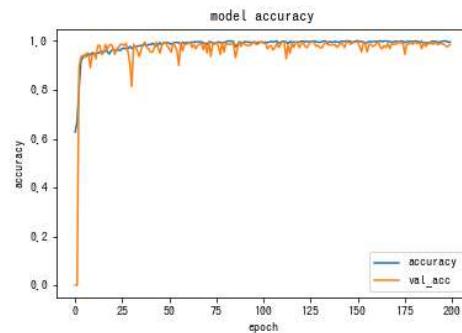


図4 識別率（データセット1）
Figure 4 Model accuracy (Dataset 1)

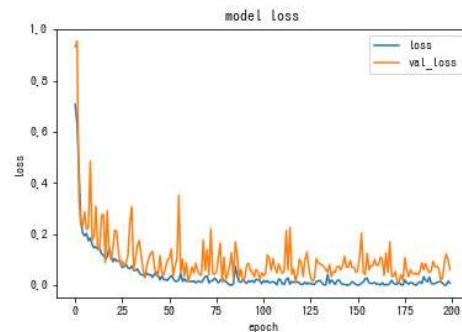


図5 損失率（データセット1）
Figure 5 Model loss (Dataset 1)

表3 データセットごとの識別率と損失値
Table 3 Accuracy and loss for each Dataset

データセット	識別率 [%]	損失値
1	99.25	0.0334
2	99.62	0.0070
3	96.60	0.0791
4	98.11	0.1029
5	99.62	0.0177
6	96.98	0.1078
7	99.62	0.0136
8	97.74	0.1148
9	99.62	0.0232
10	98.87	0.0342
平均値	98.60	0.0534

表4 誤認識リスト
Table 4 Misrecognition List

データセット	症例
1セット目	収縮期駆出性雑音
	僧帽弁閉鎖不全症
2セット目	急性前壁心筋梗塞例
	正常心音
	異常 I I 音(亢進)
	肥大型閉鎖性心筋症
	僧帽弁狭窄兼僧帽弁閉鎖不全
3セット目	収縮期駆出性雑音
	僧帽弁閉鎖不全症
	正常心音
	異常 I I 音(亢進)
	肥大型閉鎖性心筋症
4セット目	僧帽弁閉鎖不全症
	正常心音
5セット目	僧帽弁閉鎖不全症
6セット目	正常心音
	収縮早期雑音 無害性心雜音
	僧帽弁閉鎖不全症
7セット目	正常心音
8セット目	正常心音
	拡張中期雑音
	肥大型心筋症 非対称性心室中隔肥厚
9セット目	急性前壁心筋梗塞例
10セット目	正常心音
	心房細動 駆出性雑音を伴う例

2.9 考察

表4の誤認識を起こした症例リストから、複数のデータセットを通じて誤認識を起こしたものをピックアップして誤認識の原因を考察する。まずは、データセット3, 4, 6, 7, 8, 10で誤認識を起こした正常心音と、機能性心雜音である。機能性心雜音は、心音に雜音が含まれるが、病的でないものであり、識別が難しいとされているものである。また、全くの正常な心音を異常であると判別してしまった原因是、訓練データ数の少なさに原因があると考えられる。本実験で用いたデータセットに含まれる正常心音は、医学書に含まれるサンプル心音と、先行研究で実験協力者から採取された採取心音の2種類であるが、今回誤認識を起こしたものは前者のみである。正常な採取心音のデータ数は、画像にして1350枚であるが、正常なサンプル心音は50枚しかないため、学習があまり進まなかったと考えられる。先行研究では心音を5秒ごとに区切っていたため、そのままでは画像データの枚数が非常に少なく、データ拡張

を行って20倍に増やすことで対応していたが、本論文ではスライディングウィンドウ法によりある程度の枚数を確保できており、元の心音データから必要以上にサンプルを取ることは過学習に繋がると考えたため、データ拡張は行っていない。

次に、データセット1, 3, 4, 5, 6で誤認識を起こした僧帽弁閉鎖不全症である。この症例に関する誤認識はすべて単一の心音データを元に作られた画像群によるものである。以下の図6に実際に誤認識を起こしたこの症例の画像のうち1枚を示す。この症例は、収縮期雜音と、III音及びIV音が同時に聴取される「ギャロップリズム」と呼ばれる音が特徴的な症例である。しかし、ギャロップリズムは肺動脈領域で聴取される音であり、心尖部の音声データには収縮期雜音しか含まれていない。また、収縮期雜音に関しても、切り取った画像によっては含まれておらず、さらにこの心音データに含まれる外部ノイズによって、判別が難しくなっている可能性がある。この心音データの引用元である、『ぐんと身につく心音聽診技術』[15]から引用した心音データの多くは、実際の診察と比較的近い状況で心音が採取されており、多少の外部ノイズが含まれている。しかしながら、ノイズの除去を行うことで判断に必要な心雜音まで除去されてしまう可能性があるため、今回はノイズ除去処理を行っていない。より識別率を高めるためには、前処理の方法を検討する必要があると考えられる。

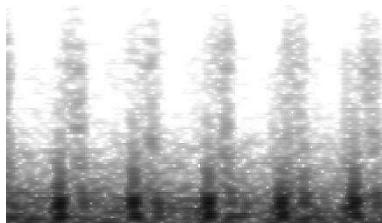


図6 僧帽弁閉鎖不全症のスペクトログラム
Figure 6 Spectrogram Image of Mitral Regurgitation

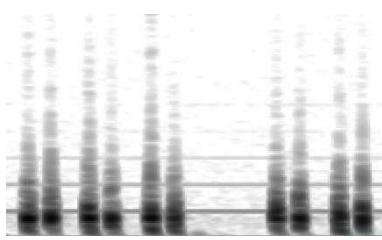


図7 心拍停止のスペクトログラム
Figure 7 Spectrogram Image of Cardiac Arrest Spectrogram

3. 置み込みニューラルネットワークとLSTMを用いた識別実験

3.1 実験概要

本実験では、スライディングウィンドウ法及び、置み込みニューラルネットワークとLSTMを組み合わせた学習モデルによる、心音の識別実験を行う。前章の置み込みニューラルネットワークのみを採用した実験と比較し、モデルの評価を行う。実験手順は前章の実験と同様である。

3.2 データセット

前章のCNNのみを用いた学習実験と比較するため、本実験では、前章でスライディングウィンドウ法を用いて製作したデータセットと同様のものを使用した。しかし、LSTMに時系列データを学習させるためには、入力データの時系列を考慮させる必要があるため、入力データが時系

列の順に整列するよう、時系列に沿って連番になるよう画像ファイル名をリネームし、名前順にソートを行った。次に、時系列を考慮しない通常の機械学習の場合、訓練データは各エポックごとにランダムにシャッフルされるが、本実験では訓練データの時系列を崩さないよう、シャッフルは行わないよう設定した。

3.3 学習モデル

学習モデルにはCNN及びLSTMを用いた。以下の表7に学習モデルの詳細を示す。置み込み層とブーリング層を含むCNN部分の学習モデルは、前章の実験と同じ構成であり、その後にLSTMを接続した。なおLSTMでは、time stepsと呼ばれる、系列データの長さを表す引数（パラメータ）が必要となるため、置み込み層やブーリング層においてもtime stepsを考慮できる、TimeDistributedラッパーを利用している。本実験では、スライディングウィンドウ法でトリミングした画像データ5枚を1つのまとまり（1サンプル）とするため、time stepsは5とした。活性化関数には、Kerasの公式ドキュメントで、リカレントニューラルネットワークに対してよいとされる[9]、RMSpropを採用した。

表5 CNN-LSTMモデルの構成
Table 5 CNN-LSTM Model

Layer	Kernel	Outputs	Function	
input		5×25×25×3		
TimeDistributed(conv1)	3×3	5×25×25×32	relu	
TimeDistributed(conv2)	3×3	5×25×25×32	relu	
TimeDistributed(pool1)	2×2	5×13×13×32	max	dropout (ratio = 0.3)
TimeDistributed(conv3)	3×3	5×13×13×64	relu	
TimeDistributed(conv4)	3×3	5×13×13×64	relu	
TimeDistributed(pool2)	2×2	5×7×7×64	max	dropout (ratio = 0.3)
TimeDistributed(conv5)	3×3	5×7×7×128	relu	
TimeDistributed(conv6)	3×3	5×7×7×128	relu	
TimeDistributed(pool3)	2×2	5×4×4×128	max	dropout (ratio = 0.3)
TimeDistributed(flatten)		5×2048		
LSTM	3×3	5×100	tanh	dropout (ratio = 0.3)
dense		5×2	softmax	

3.4 実験結果

以下の図表に実験の結果を示す。コンピュータによる学習の経緯と結果を示したものであり、図8は識別率、図9は損失値を表したグラフである。10セットの実験において、すべて同様のグラフが出力された。図8より、どのデータセットに対する実験でも、訓練データによるテストの識別率（青線）と比較して、バリデーションの識別率（オレンジ線）が向上しない事が見て取れる。また、図9より、バリデーションの損失値が大きく変動し、高い値となっていることから、学習が上手く進んでいない事が見て取れる。これらより、本実験では、コンピュータが訓練データに対して過剰に適応してしまう、過学習が起きていると考えられる。表6に学習が終了した後のテストによる識別率と損失値及びその平均値を示す。表7に誤認識リストを示す。なお、本実験では誤認識を起こした症例の種類が少なく、各セットで重複していた為、誤認された回数が多い順にリストにまとめた。チェックマークがついているデータセットで、各症例の誤認識を起こしたことを表している。表6より、本実験は前章の実験と比較して、識別率は低く、損失値は高いという、学習が出来ていないことを示す結果が見て取れる。

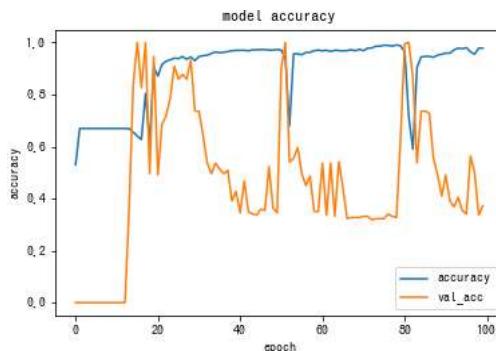


図 8 認識率 (データセット1)
Figure 8 Model accuracy (Dataset 1)

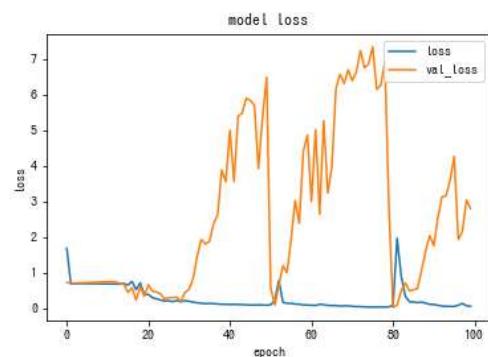


図 9 損失値 (データセット1)
Figure 9 Model loss (Dataset 1)

表 6 データセットごとの認識率と損失値
Table 7 Accuracy and loss for each Dataset

データセット	認識率 [%]	損失値
1	96.23	0.5573
2	96.23	0.4833
3	98.49	0.0964
4	96.23	0.3531
5	96.23	0.2725
6	95.47	0.5401
7	94.72	0.3398
8	96.23	0.3709
9	96.23	0.3486
10	96.23	0.4585
平均値	96.23	0.3821

表 7 誤認識リスト
Table 7 Misrecognition List

症例	データセット										回数
	1	2	3	4	5	6	7	8	9	10	
収縮期駆出性雑音	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
僧帽弁閉鎖不全症	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
正常心音 機能性雑音	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
正常心音 無害性雑音	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	8
大動脈弁狭窄	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
正常心音3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	7
僧帽弁狭窄症	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
肺血栓塞栓症	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	6
心不全	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
拡張型心筋症	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	5
正常心音2											2

3.5 考察

学習が上手く進まず、低い認識率が出た原因として、まず考えられるのはサンプル数の減少である。本実験ではLSTMに時系列を認識させるため、時系列順に並んだ5枚の画像を1つのサンプルとして学習に用いている。したがって、サンプル数は単純に5分の1となっており、これがモデルの汎化性能に悪影響を及ぼした結果、過学習を起こした可能性がある。また、表7の誤認識リストの異常心音は、すべて『ぐんと身につく心音聴診技術』を引用元とする心音データである。前述の通り、これらの心音データは医学書に添付された心音としては外部ノイズが比較的多く含まれており、医学に従事していない人間はスペクトログラム図を一見しても、心雜音を認識することは難しい。しかしながら、前章の実験ではある程度識別出来ていたことから、畳み込み層で特徴量の抽出は可能であると考えられる。

一方で、前章の実験と同様に、本提案手法のねらいである、「心音に突発的な異常が見られる症例」に対しては誤認識を起こしておらず、正しく認識をすることが出来た。この結果より、先行研究で挙げられた課題に対しては、本提案手法は有効であると考えられる。

4. おわりに

本論文では、先行研究で挙げられた、「心音に突発的な異常が現れる症状」をコンピュータが正しく認識できないという課題に対して、時系列を考慮できる学習モデル用いることで、心音の突発的な変化を認識出来るのではないかと仮説を立て、スライディングウインドウ法およびLSTMを用いた2種類の識別実験を実施し、提案手法の有効性を検討した。

第3章の実験では、スライディングウインドウ法に基づいて製作したデータセットを用いて、畳み込みニューラルネットワークによる学習と、異常心音の識別実験を行った。その結果、深層学習の学習モデル自体は先行研究とほぼ同様であるが、本論文のターゲットである期外収縮や心拍停止などの症例を正しく識別することに成功した。

第4章の実験では、スライディングウインドウ法に加え、LSTMを学習モデルに加えて識別実験を行った。時系列を考慮できる学習モデルにする事で、より確実に心音の周期性を認識させる狙いがあったが、LSTMの構造の都合上、データセットのサンプル数を犠牲にしてしまい、ターゲットとは関係のない症例に関して誤認識を起こす結果となつた。しかしながら、第3章の実験と第4章の実験の両方でターゲットの症例を正しく識別出来たことから、提案手法の有効性が示唆された。

今後の課題としては、LSTMを用いた学習モデルを上手く機能させる為、スライディングウインドウ法のオーバーラップを小さくし、再度第五章と同様の実験を行なうことが挙げられる。第3章で用いたスライディングウインドウ法とCNNの組み合わせは時系列を考慮している訳ではなく、

長期的、断続的な変化をする時系列データを正しく出来ない為、突発的な変化の頻度が落ちると、変化の無い部分の識別は不可能になると考えられる。頻度の低い期外収縮の識別や、過去の系列データを学習して将来起こる異常の予測を行うといったタスクに対して有効なLSTMを機能させることは、本研究の今後を考える上で非常に重要である。LSTMは動画像の学習に効果的であるため、オーバーラップを小さくすることで、サンプル数を増やすとともに、画像を動画像のように滑らかにするという狙いがある。

参考文献

- [1] 厚生労働省政策統括官（統計・情報政策担当），“平成30年 我が国の人団動態 —平成28年までの動向—,” 28 3 2018. [オンライン]. Available: <https://www.mhlw.go.jp/toukei/list/dl/81-1a2.pdf>. [アクセス日: 21 1 2020].
- [2] Mathworks®, “畳み込みニューラルネットワーク（CNN） - MATLAB & Simulink,” [オンライン]. Available: <https://jp.mathworks.com/discovery/convolutional-neural-network.html>. [アクセス日: 4 2 2020].
- [3] Jonathan Rubin, Rui Abreu, Anurag Ganguli, Saigopal Nelaturi, Ion Metai, Kumar Sricharan, “Recognizing Abnormal Heart Sounds Using Deep Learning,” 国際人工知能学会, 2017.
- [4] 北田聖, “深層学習を用いた心音の識別技術に関する一検討,” 2019.
- [5] 赤石誠, 香坂俊, 電子聴診器でぐんと身につく心音聴診技術, 株式会社メディカ出版, 2013.
- [6] 沢山俊民, CDによる心音シャワー・聴診の達人, 株式会社 日本医学出版, 1996.
- [7] 吉澤実, 高崎航, 大村廉, “加速度ベース行動認識におけるレスポンス時間短縮のためのパラメータ検討,” 2013.
- [8] Diederik P. Kingma, JimmyLeiBa, “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION,” 2015. [オンライン]. Available: <https://arxiv.org/pdf/1412.6980v8.pdf>. [アクセス日: 5 2 2020].
- [9] Keras Documentation, “最適化 -Keras Documentation,” [オンライン]. Available: <https://keras.io/ja/optimizers/#rmsprop>. [アクセス日: 4 2 2020].
- [10] librosa development team, “LibROSA — librosa 0.7.2 documentation,” [オンライン]. Available: <https://librosa.github.io/librosa/>. [アクセス日: 28 1 2020].