

HPSS を用いた呼吸音の自動分類

丸橋 優生^{†*}・浅谷 尚希[†]・陸 慧敏[†]・神谷 亨[†]・間普 真吾^{††}・木戸 尚治^{†††}

[†]九州工業大学 〒804-8550 福岡県北九州市戸畑区仙水町 1-1

^{††}山口大学 〒755-8611 山口県宇部市常盤台 2-16-1

^{†††}大阪大学 〒565-0871 大阪府吹田市山田丘 2-2

(受理日：2021年4月22日，採択日：2021年6月7日)

Automatic Classification of Respiratory Sounds using HPSS

Yuki MARUBASHI[†], Naoki ASATANI[†], Huimin LU[†], Tohru KAMIYA[†], Shingo MABU^{††}, and Shoji KIDO^{†††}

[†]Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-Ku, Kitakyushu, Fukuoka, 804-8550 Japan

^{††}Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi, 755-8611 Japan

^{†††}Osaka University, 2-2 Yamadaoka, Suita, Osaka, 565-0871 Japan

(Received on April 22. In final form on June 7, 2021.)

Abstract : Respiratory disease is a serious illness that accounts for three of the top ten causes of death in the world, and approximately eight million people died worldwide each year. Early detection and early treatment are important for the prevention of illness due to these diseases. Currently, auscultation is performed for the diagnosis of respiratory diseases, however there is a problem that quantitative diagnosis is difficult. Therefore, in this paper, we propose a new automatic classification method of respiratory sounds to support the diagnosis of respiratory diseases on auscultation. In the proposed method, respiratory sound data is converted into a spectrogram image by applying the short-time Fourier transform. Then, we apply HPSS (Harmonic/Percussive Sound Separation) algorithm to the respiratory sound spectrogram to separate it into a harmonic spectrogram and a percussive spectrogram. The three generated spectrograms are used for classification of respiratory sounds by CNN (Convolutional Neural Network) and SVM (Support Vector Machine) classifiers. Our proposed method obtained superior classification performance compared to the case without applying HPSS and satisfactory results are obtained.

Keywords : Respiratory Sounds, Computer Aided Diagnosis, Short-Time Fourier Transform, Harmonic/Percussive Sound Separation, Deep Learning, CNN, SVM

1. 序 論

2019年の世界保健機関(WHO: World Health Organization)の調査によると、呼吸器疾患は、世界の死因上位10位のうち3つ(3位:慢性閉塞性肺疾患, 4位:下気道感染症, 6位:気管, 気管支, 肺癌)を占める重大な疾患であり、世界中で毎年800万人近い人々が呼吸器疾患により死亡している[1]。こうした疾患による病死の予防には、早期発見, 早期治療が重要である。

呼吸器疾患を早期に発見するための診断方法の一つに、呼吸音の聴診がある。疾患者の呼吸音は、気道や肺の異常に起因する異常音(副雑音)が含まれることが多い。副雑音は、連続性ラ音と断続性ラ音に分類される。連続性ラ音は、主に気管・気管支の異常に起因する持続異常音であり、断続性ラ音は、肺の異常に起因する突発異常音である。そのため、聴取した副雑音の特徴から患者の症状を判断することができる。聴診器を用いた聴診は19世紀初頭から用いられてきた伝統的な診断方法であり、簡便、安全、安価、短時間で診断結果が得られるなどの利点から、医療機器が発達した現在においても世界中で広く採用されている[2]。しかしながら、聴診による診断には定量的な評価基準がなく、診断結果が医師の技量や経験に依存するうえ、呼吸音

は微細な要因により変化するため、正確に聴き分けることが困難な場合もある。そのため、電子聴診器により録音された呼吸音を定量的に分析・分類し、「第二の意見」として結果を出力するコンピュータ支援診断(CAD: Computer Aided Diagnosis)システム[3]の開発が必要とされる。

上記の背景から近年、呼吸音のコンピュータ支援診断システム開発における研究の多くで、ICBHI 2017 Challenge Respiratory Sound Databaseを用いた解析が行われている。これは、現在公開されている最大の呼吸音データセットとなっており、様々な録音機器、環境下で録音されている[4]。

呼吸音の自動分類の研究では、畳み込みニューラルネットワーク(CNN: Convolutional Neural Network)を用いた手法が注目されている。CNNを用いた手法では、まず次元の呼吸音信号データを、時間-周波数解析によるスペクトログラムやスカログラムなどの二次元画像データに変換を行う。次に、それら画像データをCNNに入力として与え、呼吸音を持つ深層特徴を自動抽出する。最後に、ソフトマックス関数を有する全結合層やSVM(Support Vector Machine)などの分類器を用い、呼吸音の正常と異常の自動分類を行う。こうした手法は、次元信号から手動で抽出した特徴をもとに、機械学習を用いて分類を行う過去の手法よりも、精度良く分類を行えるという研究成果

* 連絡先: 丸橋優生 (E-mail: marubashi.yuuki641@mail.kyutech.jp)

が得られている[5-7]. しかし, 異常呼吸音の特徴は不明確であることも多く, 分類精度には依然として改善の余地がある.

そこで本論文では, 各異常呼吸音がスペクトログラム上で異なった特徴の現れ方をする点に着目し, 調波打楽器音分離 (HPSS : Harmonic/Percussive Sound Separation) によってそれら異常呼吸音の特徴を分離, 明確化することで呼吸音の分類精度の向上を図る.

2. 手法

提案手法では, 呼吸音データからスペクトログラム画像を生成し, 調波音成分画像と打楽器音成分画像に分離したのち, CNNとSVMによる呼吸音の4クラス分類 (Normal : 正常呼吸音, Crackle : 断続性ラ音, Wheeze : 連続性ラ音, Both : 断続性ラ音と連続性ラ音の両方) を行う. 提案手法全体の大まかな流れを Fig.1 に示す. 以下にその詳細を示す.

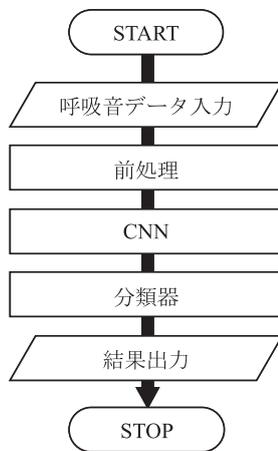


Fig.1 全体の流れ.

2.1 音声信号の前処理

ICBHI 2017 Challenge Respiratory Sound Database に収録されている呼吸音データは, サンプリングレートが異なるため, 4000 Hz にサンプリングを行う. また, 入力した各呼吸音データに対し音量の正規化を行い, データセット全体で音量をある程度統一し, 呼吸周期ごとに分割する. 最後に, 50~1800 Hz の帯域通過フィルタを呼吸音データに適用し, 聴診に重要な周波数帯域外の信号を除去する[8-10].

2.2 短時間フーリエ変換[11-13]

前処理を施した呼吸音データに対し短時間フーリエ変換を適用し, スペクトログラム画像を生成する. 以下, 短時間フーリエ変換の定義について説明する.

まず, 有限な時間長 L を有する窓関数を用意する. 本論文では, 窓関数として式(1)に示すハミング窓(Hamming Window)を用いる.

$$w(t) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi t}{L}\right), & \left(-\frac{L}{2} \leq t \leq \frac{L}{2}\right) \\ 0, & \text{(otherwise)} \end{cases} \quad (1)$$

窓関数は, 解析対象の信号を切り出す役をもつ.

時刻 t における信号の性質を調べるためには, 時刻 t を中心とした窓 $w(t)$ を信号 $f(t)$ にかけてよく, これにより信号は,

$$f_t(\tau) = w(\tau - t)f(\tau) \quad (2)$$

のように切り出される. 切り出した信号は二つの時間変数をもつ. 一つは注目している固定時刻 t , もう一つは掃引のための時刻 τ である. 窓関数で切り出された短時間区間の信号 $f_t(\tau)$ に対し, フーリエ変換を施すと,

$$F_w(\omega, t) = \int_{-\infty}^{+\infty} f_t(\tau) e^{-j\omega\tau} d\tau \quad (3)$$

となる. この式を短時間フーリエ変換(STFT : Short Time Fourier Transform)という. 短時間フーリエ変換によるパワースペクトルを式(4)に表す.

$$|F_w(\omega, t)|^2 = \left| \int_{-\infty}^{+\infty} f_t(\tau) e^{-j\omega\tau} d\tau \right|^2 \quad (4)$$

パワースペクトルを2次元平面 ($t-\omega$ 平面 : 時間-周波数平面) 上に3次的に表示すると, 時間に応じてパワースペクトルがどう変動するかを視覚的にとらえることができる. このような図をスペクトログラムと呼ぶ. 本論文では, 窓長 L を 40 ms とし, 時間位置 t を 5 ms ずつずらしながら, 短時間フーリエ変換を行う. また, スペクトログラムは, サイズが 224×224 となるようにリサイズし, 生成する.

2.3 HPSS[14-16]

多声音楽信号のスペクトログラムでは概ね, 調波音(H成分)は持続音であるため, 時間方向に滑らかな水平線として現れる. また, 打楽器音(P成分)は突発音であるため, 周波数軸方向に滑らかな垂直線として現れる. HPSSはこのようなスペクトログラムの滑らかさの異方性に着目し, スペクトログラムを水平成分と垂直成分に分離することにより各信号を分離する手法である.

本論文では, 呼吸音のうち連続性ラ音がH成分に近い特徴を持ち, 断続性ラ音がP成分に近い特徴をもつことから, この手法が異常呼吸音の特徴抽出に応用できると考え, 呼吸音スペクトログラムに対しHPSSを適用する.

HPSSにはいくつかの手法があるが, 本論文ではメディアンフィルタを用いる手法によりHPSSを行う. メディアンフィルタは信号中のあるサンプル点の信号値を, サンプル点周りの窓内の中央値で置き換える平滑化フィルタである. 入力信号を $x(n)$, メディアンフィルタのフィルタ長を l , フィルタ適用後の出力を $y(n)$ とすると, メディアンフィルタは次のように定義される.

$$y(n) = \text{median} \left\{ x(n-k) : n+k, k = \frac{l-1}{2} \right\} \quad (5)$$

ただし, l が偶数の場合は, 窓内の中央にある2つの値の平均値を中央値とする.

入力スペクトログラム O の時刻 t におけるスライスを O_t , 周波数 k におけるスライスを O_k とする. O_t 上でメディアンフィルタを適用することにより, P成分を強調したスライス P_t を生成できる.

$$P_t = M\{O_t, l_p\} \quad (6)$$

ここで, M はメディアンフィルタを表し, l_p はP成分を強調する際のフィルタ長を表す. すべての時刻についてP成分強調スライス P_t を求め, それらを組み合わせることによりP成分を強調したスペクトログラム P を得る.

同様に, H成分を強調した周波数スライス H_k は O_k 上で

メディアンフィルタを適用することにより生成できる。

$$H_k = \mathcal{M}\{O_k, l_H\} \quad (7)$$

ここで、 l_H は H 成分を強調する際のフィルタ長を表す。式(7)で生成したすべての周波数スライスを組み合わせることにより、H 成分を強調したスペクトログラム H を得る。

以上の手順で作成したスペクトログラム P および H を用い、元のスペクトログラム O に適用するソフトマスクを生成する。マスクは、ウィーナーフィルタに基づいて次のように定義する。

$$M_{H_{t,k}} = \frac{H_{t,k}^p}{H_{t,k}^p + P_{t,k}^p} \quad (8)$$

$$M_{P_{t,k}} = \frac{P_{t,k}^p}{H_{t,k}^p + P_{t,k}^p} \quad (9)$$

ここで、 p はスペクトログラムの各要素を強調する度合いを示す。通常、 p は 1 または 2 の値で設定する。

その後、H 成分、P 成分のスペクトログラムを次式により復元する。

$$\hat{H} = \hat{O} \otimes M_H \quad (10)$$

$$\hat{P} = \hat{O} \otimes M_P \quad (11)$$

ここで、 \otimes は要素ごとの乗算を表し、 \hat{O} は元の複素数値スペクトログラムを表す。

以上の手順により生成した画像の例を Fig.2~Fig.4 に示す(横軸は時間、縦軸は周波数を表す)。なお、本論文では HPSS を行う際、フィルタ長 l_H 、 l_P はともに 15 とし、ウィーナーフィルタの強調度 p は 1 とする。この値は、呼吸音を持つ特徴の欠落が生じないように、経験的に決定した。

2.4 CNN と SVM による分類

本論文では、VGG 16[17]のネットワーク構造をベースに南らが改良した CNN モデル[18]を用い、画像特徴の抽出を行う。採用したモデルの構成を Table 1 に示す。採用

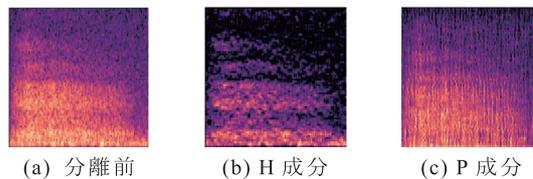


Fig.2 Normal のスペクトログラム。

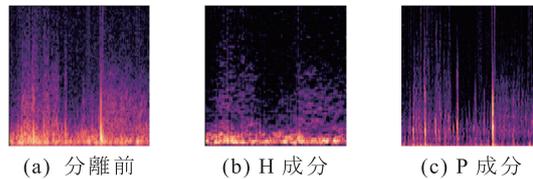


Fig.3 Crackle のスペクトログラム。

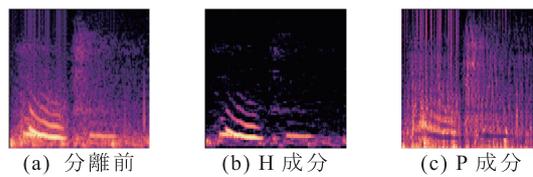


Fig.4 Wheeze のスペクトログラム。

モデルと VGG 16 との主な相違点としては、フィルタ数の削減、畳み込み層の後にバッチ正規化層を追加していること、全結合層の代わりに GAP(Global Average Pooling)層を採用していることが挙げられる。

なお、各モデルの学習時には、Keras[19]で実装されている Adam(Adaptive moment estimation)を用い、最適化を行った。事前学習は利用せず、ICBHI 2017 Challenge Respiratory Sound Database に収録されているデータを用いて学習を行った。各ハイパーパラメータは、バッチサイズを 32、最大エポック数を 50、そして学習率を 0.001 とした。

分類においては、Fig.5 に示すように、生成した 3 種類の画像を個別に CNN に入力し、すべての CNN の GAP 層出力を結合して SVM 分類器へ入力する。また、SVM については線形カーネルを用いたソフトマージン SVM(L 1 SVM)を採用し、一対一方式により呼吸音の 4 クラス分類(Normal, Crackle, Wheeze, Both)を行う。

Table 1 CNN モデルの構成。

Layer	Filter size	Stride	Output size	Remarks
Input	—	—	224×224×3	—
Conv.	3×3	1×1	224×224×16	ReLU, BN
Conv.	3×3	1×1	224×224×16	ReLU, BN
Max Pooling.	2×2	2×2	112×112×16	—
Conv.	3×3	1×1	112×112×16	ReLU, BN
Conv.	3×3	1×1	112×112×16	ReLU, BN
Max Pooling.	2×2	2×2	56×56×16	—
Conv.	3×3	1×1	56×56×32	ReLU, BN
Conv.	3×3	1×1	56×56×32	ReLU, BN
Conv.	3×3	1×1	56×56×32	ReLU, BN
Max Pooling.	2×2	2×2	28×28×32	—
Conv.	3×3	1×1	28×28×32	ReLU, BN
Conv.	3×3	1×1	28×28×32	ReLU, BN
Conv.	3×3	1×1	28×28×32	ReLU, BN
Max Pooling.	2×2	2×2	14×14×32	—
Conv.	3×3	1×1	14×14×64	ReLU, BN
Conv.	3×3	1×1	14×14×64	ReLU, BN
Conv.	3×3	1×1	14×14×64	ReLU, BN
GAP	—	—	1×64	—
FC	—	—	1×4	Softmax

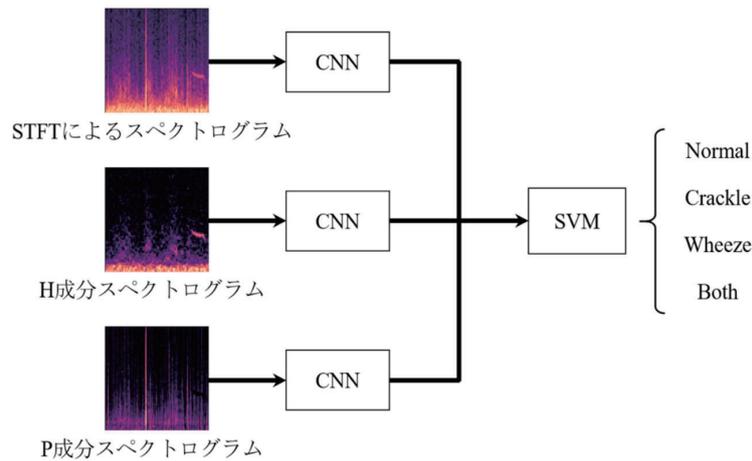


Fig.5 CNN と SVM による呼吸音の分類の流れ.

3. 実験と結果

3.1 呼吸音データの詳細

実験では, ICBHI 2017 Challenge Respiratory Sound Database に収録されている呼吸音データセットに対し, 提案手法を適用する. データセットには, 126 人の被験者から電子聴診器により聴取した 920 個の音声ファイルが含まれる. 各音声ファイルは, 呼吸周期により区切られており, それぞれ Normal, Crackle, Wheeze, Both のラベルが付与されている. 各クラスの呼吸データ数を Table 2 に示す. なお, 本論文における呼吸音の分類精度は Rocha[20]らによる分類法に基づいている.

Table 2 各呼吸音のデータ数.

Class	Number
Normal	3642
Crackle	1864
Wheeze	886
Both	506

3.2 評価方法

提案手法の評価は 5 分割交差検証法により行う. 分割した各データセットのデータ数は Table 3 に示す通りである.

また, 分類性能は ICBHI 2017 Challenge の評価指標に基づき, Average Score(AS)と Harmonic Score(HS)の 2 つの指標により評価する. それぞれの指標は, Table 4 に示すような混同行列に対し, 以下の式により算出する.

$$\text{Sensitivity(SE)} = \frac{C_c + W_w + B_b}{C + W + B} \quad (12)$$

$$\text{Specificity(SP)} = \frac{N_n}{N} \quad (13)$$

$$\text{Average Score(AS)} = \frac{\text{SE} + \text{SP}}{2} \quad (14)$$

$$\text{Harmonic Score(HS)} = \frac{2 \times \text{SE} \times \text{SP}}{\text{SE} + \text{SP}} \quad (15)$$

ここで, AS は SE と SP の平均を, HS は SE と SP の調和平均をそれぞれ表す.

Table 3 各データセットのデータ数.

	Dataset1	Dataset2	Dataset3	Dataset4	Dataset5
Normal	729	729	728	728	728
Crackle	373	373	373	373	372
Wheeze	178	177	177	177	177
Both	102	101	101	101	101

Table 4 混同行列の例.

		予測ラベル				合計
		Crackle	Wheeze	Both	Normal	
正解ラベル	Crackle	C_c	C_w	C_b	C_n	C
	Wheeze	W_c	W_w	W_b	W_n	W
	Both	B_c	B_w	B_b	B_n	B
	Normal	N_c	N_w	N_b	N_n	N

Table 5 性能評価の結果.

	SE	SP	AS	HS
HPSS 適用なし [CNN+Softmax]	0.31	0.87	0.59	0.46
HPSS 適用なし [CNN+SVM]	0.45	0.82	0.64	0.58
提案手法	0.51	0.83	0.67	0.63
Multi Image + VGG 16[18]	0.47	0.67	0.57	0.55
STFT + Wavelet, SVM[20,21]	0.20	0.78	0.47	0.32

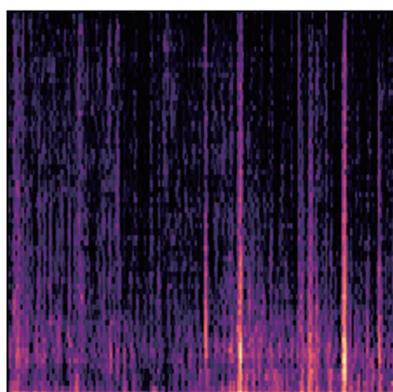
Table 6 混同行列の比較 (C : Crackle, W : Wheeze, B : Both, N : Normal).

(a) 提案手法

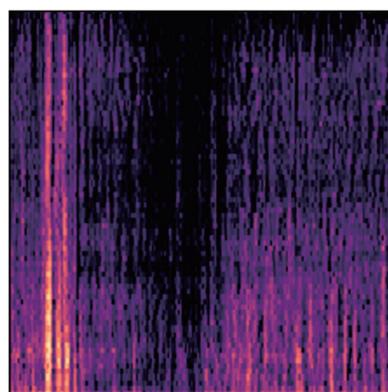
		予測ラベル			
		C	W	B	N
正解ラベル	C	999	26	45	794
	W	25	503	105	243
	B	78	173	169	86
	N	462	132	42	3006

(b) HPSS 適用無し

		予測ラベル			
		C	W	B	N
正解ラベル	C	844	26	28	966
	W	37	456	120	273
	B	91	155	151	109
	N	460	144	39	2999



(a) Normal を Crackle と誤分類



(b) Crackle を Normal と誤分類

Fig.6 誤分類されたスペクトログラムの例.

3.3 実験結果

Table 5 に提案手法の性能評価結果, Table 6 に混同行列の比較結果を示す. 表では比較のため, HPSS を適用せず, STFT によるスペクトログラム (分離前のスペクトログラム) のみを用いた CNN と Softmax 関数による分類と, CNN と SVM による分類を行った場合の結果を示している. 提案手法を, HPSS を適用しない場合と比較した結果, AS が 3.4%, HS が 5.6% それぞれ向上した. また, 本手法の有効性を示すため, 同じデータセットを用いた他の手法と比較を行う. その結果, Table 5 より従来法に比べ精度の向上が確認された.

4. 考察

HPSS による異常呼吸音の特徴分離を試みた結果, Table 5 に示す通り, すべての呼吸音クラスにおいて識別性能の向上が見られた. 特に, 断続性ラ音 (Crackle) の特徴を明確化したことにより, Crackle クラス, Both クラスの識別性能が底上げされた. 一方で, 広範囲にわたって現れる正常呼吸音 (Normal) や連続性ラ音 (Wheeze) の特徴は, 分離前のスペクトログラムでも十分に明瞭であるためか, HPSS を適用することによる大きな変化が見られないケースもあった.

提案手法を適用した場合の誤分類で特に多かったのが, Fig.6 に示す Normal と Crackle である. これは, Normal クラスの呼吸音で聴取された胸膜摩擦音, 聴診器摩擦音などの摩擦音が, Crackle と似た特徴を持っており, スペクトログラム上で両者の特徴を明確に区別できなかったために生じたものと考えられる. このような, スペクトログラム上で区別が困難な特徴を如何にして差別化するかが今後の課題となる. 断続性ラ音はそれぞれ, 特徴的なピーク周波数 (水泡音: 500 Hz 以下, 捻髪音: 1000 Hz) と持続時間 (水泡音: 10 ms 程度, 捻髪音: 5 ms 以下) を有する [10]. 一方で, 摩擦音は広い周波数帯域でピークを持つ傾向がある. そのため, 短時間フーリエ変換を行う際の窓長を変更し,

時間/周波数分解能の調整を行うことにより, それらの特徴をより明確化できる可能性がある. または, MFCC (Mel Frequency Cepstral Coefficient) 特徴量などのスペクトログラム以外の特徴量を分類器に同時入力することにより, 改善が見込めるのではないかと考えられる.

さらに, 本論文では H 成分, P 成分スペクトログラムとともに, 同じ CNN モデルを用いた特徴抽出を行ったが, 各特徴抽出に特化したネットワークモデルの改良を行うことにより, 分類性能が向上する余地があると考えられる. 一般画像に対する CNN では, フィルタサイズは正方形に設計することが一般的であるが, スペクトログラムは縦軸と, 横軸で次元が異なるため, フィルタサイズを正方形から長方形に変更することにより, 周波数軸や, 時間軸に推移する特徴を表現可能になると考えられる. 最後に, 本研究では, オリジナルの VGG 16 を実装するための計算機資源が不足するため, 改良モデルとの性能比較が困難であった. 今後は, 高性能な計算機を導入し, 比較検討を行う予定である.

5. 結論

本論文では, 呼吸音スペクトログラムに現れる異常呼吸音の特徴を, HPSS を用いて抽出したのち, CNN と SVM による呼吸音の自動分類を行うための CAD システムの開発を行った. 提案手法を HPSS 適用なしの場合と比較した結果, すべての呼吸音クラスにおいて分類精度の向上が見られ, AS は 3.4%, HS は 5.6% それぞれ向上した.

今後は分類性能のさらなる向上のため, CNN モデルの改良や, 現状のスペクトログラムでは識別困難な特徴を差別化するための, 新たな特徴量の導入を検討する必要がある.

謝辞

本研究は, 文部科学省科学研究費補助金 (21H03840) の

補助を受けている。

文 献

- [1] World Health Organization, The top 10 causes of death : <https://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
- [2] 川城丈夫, 阿部直, 菊池功次, 米丸亮, 清川浩, CDによる聴診トレーニング呼吸音編改訂第2版, 南江堂, pp.1-6, pp.25-64, 2011.
- [3] K. Doi, “Computer-aided diagnosis in medical imaging : Historical review, current status and future potential”, *Computerized Medical Imaging and Graphics*, vol.31, pp.198-211, 2007.
- [4] ICBHI 2017 Challenge Respiratory Sound Database : https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge.
- [5] K. Minami, H. Lu, H. Kim, S. Mabu, Y. Hirano, S. Kido, “Automatic Classification of Large-Scale Respiratory Sound Dataset Based on Convolutional Neural Network”, 2019 19th International Conference on Control, Automation and Systems, pp.804-807, 2019.
- [6] F. Demir, A. Sengur, V. Bajaj, “Convolutional Neural Network based efficient approach for classification of lung diseases”, *Health Information Science and Systems*, vol.8, no.4, 2020.
- [7] F. Demir, A. M. Ismael, A. Sengur, “Classification of Lung Sounds With CNN Model Using Parallel Pooling Structure”, *IEEE Access*, vol.8, pp.105376-105383, 2020.
- [8] 竹村直行, “肺の聴診呼吸音の分類を中心に”, *動物の循環器*, vol.30, pp.21-26, 1997.
- [9] 八田健司, 多田昌司, 中村英夫, 吉田正樹, “独立成分分析法を用いた心音呼吸音分離システムの開発”, *電子情報通信学会, 信学技報*, vol.103, no.731, pp.39-44, 2004.
- [10] 長坂行雄, 土谷美知子, 坂口才他, “臨床で役立つ肺音の聴診”, *洛和会病院医学雑誌*, vol.29, pp.1-7, 2018.
- [11] 和田成夫, よくわかる信号処理, 森北出版, pp.73-78, 2009.
- [12] L.コーエン著, 吉川昭, 佐藤俊輔訳, 時間-周波数解析, 朝倉書店, pp.93-112, 2000.
- [13] 小野順貴, “短時間フーリエ変換の基礎と応用”, *日本音響学会誌*, vol.72, no.12, pp.764-769, 2016.
- [14] N. Ono, K. Miyamoto, H. Kameoka, J. L. Roux, “Harmonic and Percussive Sound Separation and Its Application to MIR-Related Tasks”, *Springer* 274, pp.213-236, 2010.
- [15] 橘秀幸, 小野順貴, 嵯峨山茂樹, “スペクトルの時間変化に基づく: 音楽音響信号からの歌声成分の強調と抑圧”, *情報処理学会研究報告*, vol.2009-MUS-81, no.12, pp.1-6, 2009.
- [16] D. Fitzgerald, “Harmonic/Percussive Separation using Median Filtering”, 13 th International Conference on Digital Audio Effects (DAFx-10), 2010.
- [17] K. Simonyan, A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, arXiv preprint arXiv : 1409.1556, 2014.
- [18] K. Minami, H. Lu, H. Kim, S. Mabu, “Automatic Classification of Respiratory Sounds Based on Convolutional Neural Network with Multi Images”, 2020 5th International Conference on Biomedical Imaging Signal Processing, 2020.
- [19] Keras Documentation : <https://keras.io/ja/>.
- [20] B. M. Rocha, D. Filos, L. Mendes et al, “An open access database for the evaluation of respiratory sound classification algorithms”, *Physiological Measurement*, vol.40, no.3, 035001, 2019.
- [21] G. Serbes, S. Ulukaya, Y. P. Kahya, “An Automated Lung Sound Preprocessing and Classification System Based on Spectral Analysis Methods”, *Precision Medicine Powered by pHealth and Connected Health*, pp.45-49, 2018.