

一般口演

一般口演13

画像情報・生体信号処理

2022年11月19日(土) 09:00 ~ 10:45 I会場 (207会議室)

[3-I-1-02] アラーム音自動検出システムに適したマイクロホン下限性能の検討

*岸本 和昌¹、竹村 匡正²、山本 豪志朗^{1,3,4}、杉山 治⁵、小島 諒介³、黒田 知宏^{1,3,4}（1. 京都大学医学部附属病院, 2. 兵庫県立大学大学院応用情報科学研究科, 3. 京都大学大学院医学研究科, 4. 京都大学大学院情報学研究科, 5. 近畿大学情報学部）

*kazumasa Kishimoto¹, Tadamasa Takemura², Goshiro Yamamoto^{1,3,4}, Osamu Sugiyama⁵, Ryosuke Kojima³, Tomohiro Kuroda^{1,3,4}（1. 京都大学医学部附属病院, 2. 兵庫県立大学大学院応用情報科学研究科, 3. 京都大学大学院医学研究科, 4. 京都大学大学院情報学研究科, 5. 近畿大学情報学部）

キーワード：Clinical Alarms, Deep Learning, Sound Event Detection

医療従事者が患者の状態を把握することは重要であり、異常状態を即座に知ることによって安全性の高い医療が提供できる。医療機器はJIS規格に基づき異常状態を通知するためにアラーム音を発報するが、医療従事者は遠方や夜間の見回りのために病室で鳴るアラーム音に気付けないことがある。一方でIHE:PCDに準拠した医療機器の増加によってHIS連携が進んでいるが、ネットワーク接続非対応の小型機器も病棟に存在するため、医療機器の異常時に鳴るアラーム音を検出できれば効果的である。これまでマルチラベルを付与した重複音の識別や、エッジデバイスによる音声認識も試みられている。これらからアラーム音の機種を分類し、既にHISと連携している医療機器と区別することによって、ネットワーク接続非対応の医療機器も含めて簡易にHISと連動できる。また、多数の病室があるため患者ベッドサイドに設置したエッジデバイスを用いて検出結果を通知することで、システム全体の合理化を図ることが期待できる。我々はアラーム音自動検出システムをシングルボードコンピュータとマイクロホンのコンポーネントで検討してきたが、ベッドサイドで利用する場合は小型なパッケージが望ましい。そこでアラーム音は単純な正弦波信号であるため、より小型なマイクロホンが利用できる可能性がある。本研究では、深層学習を用いたアラーム音自動検出システムの構築を目的として、代表的ないくつかのモノラルマイクロホンを用いてアラーム音の分類性能を比較した。様々なマイクロホンを調査した結果から、アラーム音自動検出システムに用いるエッジデバイスの小型化が期待できた。

アラーム音自動検出システムに適したマイクロホン下限性能の検討

岸本 和昌^{*1}、竹村 匡正^{*2}、山本 豪志朗^{*1,3,4}、杉山 治^{*5}、小島 諒介^{*3}、黒田 知宏^{*1,3,4}

^{*1} 京都大学医学部附属病院、^{*2} 兵庫県立大学大学院応用情報科学研究科、

^{*3} 京都大学大学院医学研究科、^{*4} 京都大学大学院情報科学研究科、^{*5} 近畿大学情報学部

Investigation of Lower Limit Performance of Microphones for Automatic Alarm Sound Detection System

Kazumasa Kishimoto^{*1}, Tadamasa Takemura^{*2}, Goshiro Yamamoto^{*1,3,4}, Osamu Sugiyama^{*5}, Ryosuke Kojima^{*3}, Tomohiro Kuroda^{*1,3,4}

^{*1} Kyoto University Hospital, ^{*2} Graduate School of Applied Informatics, University of Hyogo,

^{*3} Graduate School of Medicine, Kyoto University, ^{*4} Graduate School of Informatics, Kyoto University,

^{*5} Faculty of Informatics, Kindai University

The medical staff needs to know the abnormal states of medical devices, and immediate notification of the abnormal state will enhance safer patients' medical care. The increasing number of medical devices compliant with IHE:PCD has led to their integration with hospital information systems. However, some small medical devices are still not network-compatible, and alarm sound detection with microphones on light-weight IoT devices could be one of the practical solutions for detecting their abnormal states. In this study, we propose an alarm sound detection system with a single board computer and a microphone. Alarm sound detection with a smaller microphone is appropriate since the alarm sound is a simple sinusoidal signal and it would be used at the patients' bed-side. In this study, we compared the classification performance of alarm sounds using several typical monaural microphones to develop an alarm sound detection system using deep learning. As a result, comparing various microphones, the performance of the smaller MEMS microphone was comparable to other types of microphones.

Keywords: Clinical Alarms, Deep Learning, Sound Event Detection.

1. 結論

医療従事者が患者の状態を把握することは重要であり、異常状態を即座に知ることで安全性の高い医療を提供することが期待できる。患者の治療に用いる医療機器は異常を検知すると発報する仕組みとなっており、機器側の異常状態だけでなく、間接的に患者側の異常も含まれている。このため医療機器のアラームへの対応は、治療を継続させるためだけでなく、患者の異常状態への早期の対処に繋がる。医療機器は JIS 規格に基づき異常状態を通知するためにアラーム音を発報するが、医療従事者は遠方や夜間の見回りのために病室で鳴るアラーム音に気付けないことがあり、アクシデントの報告が挙がっている¹⁾。

IHE:PCD に準拠した医療機器の増加によって病院情報システムへの連携が進んでいるが、ネットワーク接続非対応の小型機器も病棟には存在するため病院情報システムとは連携する手段がない。一方で、医療機器は異常時にアラーム音を発報するため、機種ごとに検出できれば医療安全の質の向上に寄与できる。

医療機器のアラーム音を検知する手法として、これまでマルチチャネルを付与した実生活の環境音における重複音の識別や、エッジデバイスによる音声認識も試みられている^{2),3)}。アラーム音の機種を分類し、既に病院情報システムと連携している医療機器と区別することによって、ネットワーク接続非対応の医療機器も含めて簡単に病院情報システムと連動できる。多数の病室があることで、1地点からの音声認識でどの医療機器がアラーム音を発しているかを検知することは困難である。この課題は、患者ベッドサイドにエッジデバイスを設置し検出結果を通知することで解決することが期待できる。我々はアラーム音自動検出システムをシングルボードコンピュータとマイクロホンのコンポーネントで検討してきたが、ベッドサイドで利用する場合は小型なパッケージが望ましい^{4,5)}。

アラーム音は単純な正弦波信号であるため、音声録音を記録する高性能なマイクではなく、より小型なマイクロホンで十分な音声認識精度が得られる可能性がある。

2. 目的

本研究では、深層学習を用いたアラーム音自動検出システムの構築を目的として、代表的ないくつかのモノラルマイクロホンを用いてアラーム音の分類性能を比較する。

3. 方法

事前に機器とマイクロホンとの距離と音圧を測定しながら録音したアラーム音を用いてシミュレーションデータを作成した。次に検証対象の各種マイクロホンとスピーカを、事前に録音した際の距離と音圧が同等になるよう調整し評価データセットを作成した。この評価データセットから各判別器で分類性能を評価し、アラーム音の検出におけるマイクロホンを検討した。

3.1 対象

3.1.1 マイクロホン

マイクロホンは様々な種類があるなかで、代表的な 4 種類であるダイナミック型、コンデンサ型、ECM(Electret Condenser Microphone)型、MEMS(Micro Electro Mechanical System)型を比較した(表 1)。ダイナミック型はハンドマイク、コンデンサ型は会議用マイク、ECM 型は USB 直結の小型マイク、MEMS 型は Raspberry Pi 3B+で用いる基盤モジュールである。

3.1.2 アラーム音

病院の一般病棟における医療機器の使用数が多い人工呼吸器装着患者を想定し、分類するアラーム音は、輸液ポンプ、シリンジポンプ、経腸栄養ポンプ、フットポンプ、胸腔ドレナージ、患者監視モニタ、人工呼吸器の 7 機種とした。病院

表 1 マイクロホンの仕様と録音条件

種類	メーカー/型式	周波数特性[Hz]	指向性	録音音量
ダイナミック	オーディオテクニカ/ ATR2100	50~15k	単一	0.5 ^a
コンデンサ	サンワサプライ/MM-MCU05BK	20~20k	無	0.5 ^a
ECM	Adafruit/ 3367	100~16k	無	0.5 ^a
MEMS	Knowles/SPH0645	20~20k	無	61 ^b

^a Audacity^b AlsaMixer

の個室で実機を動作させてベッドの頭側に設置したコンデンサ型マイクロホンで静音化にて録音した。この時のマイクロホンと医療機器の距離と、A 特性の騒音計で音圧を測定した。これらデータのアノテーションとスペクトル分析は Audacity¹ を用いてイベントのラベル付けと基本周波数の抽出を行った(表 2)。録音したアラーム音はピーク周波数のコントラストが明瞭な正弦波で、シリンジポンプ、経腸栄養ポンプ、フットポンプはパルス音、輸液ポンプ、胸腔ドレナージ、患者モニタ、人工呼吸器はバースト音であった。

3.2 特徴抽出

音響識別の多くはスペクトルに人の周波数知覚特性を考慮したメルフィルタバンクをかけたログメルスペクトログラムが用いられる。音響解析ライブラリは Librosa 0.9.1 を用いて、0.43 秒の入力した音響データを window size を 1024、hop length を 512、Hamming window の短時間フーリエ変換でパワースペクトログラムに変換した。次に 40 個のメルフィルタで内積を算出し、対数化して 40×40 のログメルスペクトログラムを特徴量とした。

3.3 分類器

分類器はいくつかの代表的なネットワークモデルを用いて作成した。環境構築として、プログラミング言語は Python 3.8、深層学習ライブラリは Tensorflow 2.6.1 で学習済みモデルを作成した。学習データセットはデータ拡張処理としてホワイトノイズを信号対雑音比 30dB から 0dB の範囲で 30 倍に水増しし、さらに周波数方向と時間方向に対してマスク処理を行う SpecAugment を適用した。次に重複アラーム音を分類するために、出力層は分類数の出力ノードに活性化関数としてシグモイド関数で出力するマルチラベル分類の設計とした。最後に、学習条件は 5 分割交差検証、最適化関数 Rectified Adam、エポック数 60 で学習曲線から過学習に陥っていないことを確認した。

3.3.1 CNN モデル

音響データはログメルスペクトログラムによる画像問題として分類できるため、従来から畳み込みニューラルネットワーク(以下、CNN)が用いられる。CNN のベースモデルとして分類性能と処理速度の両立を狙って設計された EfficientNetB0~B3(以下、EN-B0~3)をそれぞれ適用した⁶⁾。EfficientNet はモデル番号が大きくなるほど深いネットワーク構造を取る。ネットワーク設計として Global Average Pooling 2D 層と全結合層を追加した。

¹ <https://www.audacityteam.org/>

表 3 各マイクロホンにおける分類モデルごとの F 値

ノイズ	モデル	ダイナミック	コンデンサ	ECM	MEMS
バブル	CRNN	0.828	0.788	0.820	0.832
	EN-B0	0.901	0.873	0.897	0.898
	EN-B1	0.898	0.873	0.894	0.896
	EN-B2	0.899	0.876	0.892	0.895
	EN-B3	0.899	0.873	0.895	0.900
病棟	ViT	0.881	0.859	0.884	0.884
	CRNN	0.887	0.874	0.870	0.896
	EN-B0	0.932	0.926	0.927	0.939
	EN-B1	0.935	0.926	0.928	0.939
	EN-B2	0.934	0.925	0.929	0.941
	EN-B3	0.934	0.926	0.928	0.939
	ViT	0.932	0.920	0.923	0.936

3.3.2 CRNN モデル

音響データは時系列データであるため再帰型ニューラルネットワーク(以下、RNN)が適応できる。畳み込み再帰型ニューラルネットワーク(以下、CRNN)はスペクトログラムの周波数方向の局所的な情報を CNN で捉え、大域的な時間変化には後段の RNN によって対応する。ネットワーク構成はフィルタ数 128 の CNN 3 層に続いて、32 ユニットの双方向性 RNN 2 層の構成とした⁷⁾。

3.3.3 Vision Transformer モデル

Vision Transformer(以下、ViT)は、Attention 機構の概念を基にした Multi-head Attention から構成される Transformer のエンコーダ部を応用したネットワークモデルである⁸⁾。Transformer は自然言語処理分野から発展したが、ViT は画像を小さいパッチに分割して位置情報を埋め込むことで、各パッチを単語のように扱って推論する。ViT のネットワーク設計はパッチサイズ 4 で分割し、Multi-head Attention はヘッド数 4 の 8 層、最後に全結合層のノード数を 128 に続き 32 とした。

3.4 評価

3.4.1 シミュレーションデータ

検証に用いる重複アラーム音は、静音下で 7 機種の実機でアラーム音を鳴らし合わせて録音した音響データに 2 種類の環境音毎に信号対雑音比 0dB で重畳させたシミュレーションデータを作成した。1つ目の環境音は群衆の多重音声からなるバブルノイズとした。2つ目は一般病棟で録音したノイズデータを用いた。これには会話、足音、戸棚の開閉音、口腔内吸引、人工呼吸器の呼吸音が含まれているが、他の医療機器のアラーム音は含まれていない。

3.4.2 マイクロホンの評価データセット

マイクロホンの評価には 3.4.1 項のシミュレーションデータをスピーカ(Yamaha: YVC-1000)から鳴らして各マイクロホンで録音した。音圧が最も大きい人工呼吸器と同等の条件に合わせるよう、スピーカの出力音量は離隔距離 90cm で音圧 79.6dB と近似するように調整した。スピーカと各マイクロホン

の向きは正対となるように配置し、同様の離隔距離で録音した。この録音した音響データから Audacity を用いてイベントをラベル付けし、評価データセットとした。

3.4.3 評価手法

分類器の出力値はクラス毎にシグモイド関数から 0~1 の推定値を得るため、閾値 0.5 で 2 値分類した。音響イベントの評価ライブラリは sed_eval 0.2.1 を利用し、評価指標として F 値を算出した。5 分割交差検証で作成した 5 つの分類器から得られる F 値を相加平均した値から検討した。

4. 結果

表 3 に評価データセットで分類した F 値の結果を示し、各モデルのうち最も高値を太字とした。

病棟ノイズよりバブルノイズの方が低い結果となり、各ノイズにおいて CRNN モデルは他のモデルより F 値が 0.05 以上低かった。バブルノイズにおいては、CRNN モデル以外ではマイクロホンによる違いは F 値の差が 0.05 未満で、ViT モデルのコンデンサ型を除くと差が 0.03 未満であった。病棟ノイズにおいては、すべてのモデルで MEMS 型が最も高値であったが、CRNN モデル以外では F 値の差が 0.03 未満であった。

5. 考察

判別器は CRNN モデル以外のモデルにおいて F 値が高く、マイクロホンにおいては MEMS 型が高かった。一部の条件においてはダイナミック型がわずかに MEMS 型より上回っていたものの、アラーム音の判別においてマイクロホンの種類によって大きな差はないと考えられた。ECM 型と MEMS 型は小型な基盤モジュールであるため、マイクロホンを内蔵することによってエッジデバイスを小型なパッケージとして構築することができる。

本実験の限界として、スピーカを用いてシミュレーションデータを鳴らした録音データで評価した事と、マイクロホンと対象機器の方向と距離が一定であるため、実際とは異なる可能性がある。

6. 結論

患者ベッドサイドに設置したエッジデバイスを用いて医療機器のアラーム音から異常を検出し、ネットワークに接続されていない医療機器も病院情報システムと連携することによっ

て、安全性の高い医療を提供することが期待できる。このエッジデバイスは小型なパッケージが望まれるため、マイクロホンは小型かつ判別性能が担保されるモジュールが望ましい。

本実験において、MEMS 型は多くの条件で F 値が高値かつ、他の方式のマイクロホンと遜色ない判別性能と考えられた。今後は、MEMS 型マイクロホンを利用したエッジデバイスを構築しつつ、実行可能性を検討する必要がある。

参考文献

- 1) 石川雅彦, 斉藤奈緒美. 医療機器のアラームに関わるインシデント・アクシデント事例からみる再発防止の検討. 医療機器学. 2017; 87(3): 285-91.
- 2) Cakir E, Heittola T, Huttunen H, Virtanen T. Polyphonic sound event detection using multi label deep neural networks. 2015 International Joint Conference on Neural Networks (IJCNN). 2015; 1-7.
- 3) Cerutti G, Andri R, Cavigelli L, Farella E, Magno M, Benini L. Sound event detection with binary neural networks on tightly power-constrained IoT devices. Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design. Association for Computing Machinery. 2020; 19-24.
- 4) 岸本和昌, 竹村匡正, 杉山治, 小島諒介, 八上全弘, 南部雅幸, ほか. 深層ニューラルネットワークを用いた重複アラーム音識別システムの検討. 生体医工学 2022; 60(1): 8-15.
- 5) 岸本和昌, 竹村匡正, 杉山治, 小島諒介, 山本豪志朗, 黒田知宏. エッジデバイス上における医療機器アラーム音分類器の実行可能性の検討. 第 66 回システム制御情報学会研究発表講演会 2022.
- 6) Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. International conference on machine learning 2019; 6105-6114.
- 7) Adavanne S, Virtanen T. A report on sound event detection with different binaural features. Detection and Classification of Acoustic Scenes and Events 2017; Available at: <https://arxiv.org/abs/1710.02997v1>
- 8) Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv 2020. Available at: <http://arxiv.org/abs/2010.11929>

表 2 アラーム音の詳細

機種	メーカー / 型式	ピーク周波数 (Hz) [°]	発報時間(s)	休止時間(s)	マイクとの 距離(cm)	音量 設定	音圧(dB)
輸液ポンプ	JMS / OT-818G	(856 856 856 - 856 856) 2 回	3.34	2.99	75	6	71.9
シリンジポンプ	Terumo / TE-351	4001	0.26	0.20	60	b-2	77.7
経腸栄養ポンプ	Fresenius Kabi / APPLIX Smart	4097	0.80	0.80	64	-	64.4
フットポンプ	Covidien / SCD700	2108	0.20	1.30	245	-	61.2
胸腔ドレナージ	Medela / THOPAZ	2632	1.07	8.95	100	-	73.7
患者モニタ	NihonKoden / PVM-4761	(783 994 1181 - 1181 1569) 2 回	2.93	4.09	120	9/13	62.1
人工呼吸器	Hamilton / C1	491 828 662	1.10	4.98	90	5	79.6

° 基本周波数