

令和元年6月24日現在

機関番号：12103

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K16460

研究課題名(和文) 深層学習を用いた環境音認識による聴覚障害者支援システム基盤の開発

研究課題名(英文) Development of Hearing Impaired Support System Platform by Environmental Sound Recognition using Deep Learning

研究代表者

白石 優旗 (Shiraishi, Yuhki)

筑波技術大学・産業技術学部・准教授

研究者番号：00389214

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究は、環境音を非常に高い精度で識別しユーザに伝達する、聴覚障害者支援システム基盤の開発を目指すものである。結果、1)歩道などの騒音の多い環境下においてもスマートフォンのみで警告音5種(車のクラクション、救急車のサイレン、自転車のベル、火災報知器、騒音)を識別し、2)ユーザが選択した警告音種のみ通知可能であり、3)新規の警告音種についてはデータを追加し再学習することで識別可能となるシステムを構築した。

研究成果の学術的意義や社会的意義

1)環境音には多種のノイズが含まれ、その中から種々の移動体が発する特定の音を高精度で識別することは一般に困難な課題とみなされているにも関わらず、本研究において高い精度を示すことができたこと、並びに、2)昨今の音声認識研究の進展は目覚ましいものの、福祉工学的視点からの環境音認識研究の不足が世界的に叫ばれていることから、本研究の学術的意義を確認できる。また、本システムにより聴覚障害者や聴力の低下した高齢者等の安全性を確保することが可能になることから、社会的意義についても確認できる。

研究成果の概要(英文)：This study aims to develop a hearing impaired support system platform that identifies environmental sound with very high accuracy and transmits it to the user. As a result, 1) the developed system classified five kinds of alarm sounds (car horn, ambulance siren, bicycle bell, fire alarm, and noise) even in noisy environment such as sidewalk using a smartphone, 2) the system can notify only alarm sound types you selected, and 3) new alarm sound types can be classified by relearning using new added data.

研究分野：福祉工学 ニューラルネットワーク システム工学 データ工学 Web情報システム

キーワード：聴覚障害者支援 福祉工学 支援機器 音認識 機械学習

## 様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

### 1. 研究開始当初の背景

難聴高齢者を含む聴覚障害者が、質の高い生活(QOL)を送るためには、様々な環境音の中でも特に安全・安心な生活に直結する警告音(クラクション、救急車等)を確実に認識可能なことが強く求められる。そのため、環境音の中から特定の警告音を識別し、ユーザに伝達するシステムが必要とされている。

警告音識別には非常に高い精度が求められるのに対し、これまでに、パワースペクトル、自己相関関数、パルスニューラルネットワーク等を用いた複数の研究が報告されているものの、雑音環境下での識別精度が最大で90%程度と不十分であった。また、警告音を発する対象物の移動や周辺環境の変化による警告音の変化への対応が困難といった課題が残されていた。

一方で、深層学習と呼ばれる、認識したい実際の警告音データをコンピュータに提示することで自動的に学習し識別可能とする手法が様々な分野に適用され、大幅に識別性能を向上し、近年特に注目されている。なお、深層学習には、従来からある階層型ニューラルネットワークを多段に拡張したものが使用されている。

### 2. 研究の目的

本研究では、近年様々な分野で大幅に識別性能を向上している深層学習と、多くの人々がインターネットを通して容易に協力可能なクラウドソーシングのしくみを用いることで、環境音を非常に高い精度で識別しユーザに伝達する、聴覚障害者支援システム基盤を開発する。

その際、(1)スマートフォンやウェアラブル端末を活用することで日常的に利用可能、(2)クラウドソーシング技術により深層学習に必要な大量データを容易に確保可能、(3)学習機能により新規の環境音に対応可能、という利点を持つシステムの開発を目指す。

本助成期間においては、環境音の中でも安全・安心な生活に直結する警告音(クラクション、救急車等)を主な対象とし、日常環境の雑音下において99%以上の識別性能を目指す。更に、その他の環境音にも柔軟に対応可能なシステム基盤を開発する。

### 3. 研究の方法

研究方法は、(1)深層学習による高精度の識別性能の実現、(2)深層学習に用いる多量の学習データ採取のためのクラウドソーシング基盤の構築、(3)聴覚障害者が日常利用可能なインターフェイスシステムの作成、の3本の柱からなる。

具体的には、(1)識別対象とする環境音の選定と環境音データの採取、(2)識別アルゴリズムの検討と学習用データベース(DB)の作成、(3)識別プログラムの実装と情報伝達システムの検討、(4)情報伝達システムの実装と評価実験、(5)クラウドソーシングによる学習データ採取システムの作成と運用、の各項目を実施する。

### 4. 研究成果

(1) 識別対象とする環境音の選定と環境音データの採取に関しては、先行研究で集めていた音データ(救急車のサイレン、自転車のベル)に加え、Webページよりクラクションの音データ計18種類をダウンロードし、また火災警報音を避難訓練の際に録音することで採取した。また、識別対象音以外の音が発生した場合に対応するため、雑音として、足音、車の走行音、声、ドアの開閉音、机を叩く音、ビニール袋の擦れる音の6種を収集した。更に、Webページより車のクラクション音151種、救急車のサイレン音120種、自転車のベル103種を追加でダウンロードし、学習・評価用DBを作成した。

(2) 警告音識別のためのアルゴリズムとして、連続的に環境音を収集、閾値以上の音量データを検知した場合一定時間の音データを記録、記録された音データに対して警告種を識別、の3つのステップが必要になる。また、警告音はその性質上、単調で繰り返される傾向が強いことから、上記の閾値処理により採取された音データに対して短時間フーリエ変換(Short-Time Fourier Transform, STFT)により、パワースペクトルに変換し、更にlogスケールに変換したものをDNN(deep neural network)への入力とした。その際、DNNに雑音クラスに対応する出力細胞を追加することで、識別対象音以外の音が発生した場合に対応するアルゴリズムを開発した。

(3) 実環境においては、対象音は時間的に連続して鳴り続けるため、1回の対象音の発生に対して識別結果が複数回表示されることになる。したがって、雑音以外の警告音の重要度も考慮し、識別率と可読性を向上させるため、1回以上~10回以下の音の連続評価結果に対して、雑音を除いた中から最大のものを最終識別結果として判定するアルゴリズム(統合判定アルゴリズム)を開発した。

(4) プログラミング言語 Swift を用いて、iPhone で環境音を録音・識別並びにユーザへ通知可能なアプリケーションを開発した。その際、Apple の公開している API である BNNS (Basic neural network subroutines) を用いることで、iPhone 上で NN (neural network) を効率的に構築した。その際、3 層 NN、4 層 DNN、5 層 DNN のうちいずれかを選択可能とした。識別アプリケーションの画面を図 1 に示す。

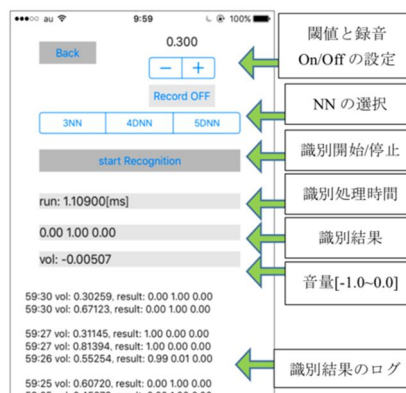


図 1 iPhone アプリの画面

(5) 識別可能な警告音にこれまでの救急車のサイレンと自転車のベルに加えて、自動車のクラクションと火災警報音を追加し、更に、雑音を第 5 クラスの対象音として学習することで雑音への対策を行なった結果、騒音の多い実環境下においても、クラクション、自転車のベル、救急車のサイレン、火災警報音全ての場合で 99%以上の識別精度を確認できた。また、未知のクラクション音に対しても 95%以上の識別率が得られた。一方で、警告音と同時に大きな雑音が発生した場合は、警告音が雑音として識別されることがあったものの、更なる統合処理を加えることで正しく警告音を伝達できる可能性が示された。

(6) 実用化に必要な課題として、学習データの更なる追加収集、識別処理と伝達方法の改善、音源方向の推定、について検討を重ねた結果、多種多様な警告音への対応が最優先と判断し、を重点的に実施することとした。具体的には、Urban Sound Dataset (世界中のユーザがアップロードした音データから都市の日常音を選択しラベリングしたデータ・セット) から車のクラクション 428 個、救急車のサイレン 929 個、自転車のベル 169 個を追加収集した。その後、それらからノイズの比較的小さいデータとして、車のクラクション 151 種、救急車のサイレン 120 種、自転車のベル 103 種 (各種 1~2 個) を選定し、これまでの識別対象 5 種 (車のクラクション、救急車のサイレン、自転車のベル、火災報知器、騒音) に追加の上、実験を行った。5-fold 法で評価した結果、約 97%の識別率を得た。すなわち、新種の音データに対して、同一のアルゴリズムにより、学習データを追加するのみで対応可能なことを確認した。

(7) 追加の学習データ採取システムについては、非営利のマイクロボランティア・クラウドソーシングプラットフォームである Crowd4U を用いることで実現可能となることを確認した。

(8) 研究期間全体を通じて実施した研究の成果をまとめると、歩道などの騒音の多い環境下においてもスマートフォンのみで警告音 5 種 (車のクラクション、救急車のサイレン、自転車のベル、火災報知器、騒音) を識別し、ユーザが選択した警告音種のみ通知可能であり、新規の警告音種についてはデータを追加し再学習することで識別可能となるシステムを構築した。今後については、これまでの成果を発展させることで、クラウドソーシングによる学習データ採取システムの実装、識別対象の拡張、識別処理と伝達方法のさらなる改善、音源方向の推定等に取り組んで行く予定である。

## 5. 主な発表論文等

[雑誌論文](計 8 件)

Shitara A., Namatame M., Shiraishi Y., Tactile Stimulus Start System Proposal for Deaf and Hard of Hearing Sprinters, Proc. 7th International Conference for Universal Design (UD 2019), 査読有, 2019, pp.1-8

Kumai K., Matsubara M., Shiraishi Y., Wakatsuki D., Zhang J., Kitagawa H., and Morishima A. Skill-and-Stress-Aware Assignment of Crowd-Worker Groups to Task Streams, Proc. the 6th AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2018), 査読有, 2018, pp.88-97

Hashimoto H., Matsubara M., Shiraishi Y., Wakatsuki D., Zhang J., and Morishima A., A Task Assignment Method Considering Inclusiveness and Activity Degree, Proc. The Second IEEE Workshop on Human-in-the-loop Methods and Human Machine Collaboration in BigData (HMData2018), 査読有, 2018, pp.3497-3502

Shitara A., Namatame M., Shiraishi Y., Proposal of a Vibration Stimulus Start System for the Deaf and Hard of Hearing, Journal on Technology & Persons with Disabilities, 査読有, 2018, pp.140-148

Shiraishi Y., Zhang J., Wakatsuki D., Kumai K., Morishima A., Crowdsourced real-time

captioning of sign language by deaf and hard-of-hearing people, International Journal of Pervasive Computing and Communications, 査読有, 2017, pp.2-25  
DOI:10.1108/IJPCC-02-2017-0014

Shionome T., Hashimoto H., Zhang J., Shiraishi Y., Wakatsuki D., Seki Y., and Morishima A., Complement of Incomplete Task Results for Real-time Crowdsourcing Interpretation, Proc. the 21th International Conference on Asian Language Processing (IALP 2017), 査読有, 2017, pp.359-362

Kumai K., Zhang J., Shiraishi Y., Wakatsuki D., Kitagawa G., 査読有, Group rotation management in real-time crowdsourcing, Proc. the 19th International Conference on Information Integration and Web-based Applications & Services (iiWAS 2017), 査読有, 2017, pp.23-31

Zhang, J., Shiraishi, Y., Kumai, K., and Morishima, A., Real-Time Captioning of Sign Language by Groups of Deaf and Hard-of-Hearing People, Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services (iiWAS 2016), 査読有, 2016, pp.54-63  
DOI:10.1145/3011141.3011143

〔学会発表〕(計9件)

土屋智彦、白石優旗、深層学習を用いたセンサグローブによる指文字認識の改良、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第9回研究会、2019

設楽明寿、生田目美紀、白石優旗、聴覚障害者へのスタート合図に最適な触覚刺激インターフェースの特定-陸上競技短距離走スタートシステムのユニバーサルザブインを目指して-、情報処理学会第81回全国大会、2019

矢野和希、白石優旗、スマートフォンを用いた深層学習による警告音認識システムの改良、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第6回研究会、2018

小松周生、白石優旗、深層学習を用いたセンサグローブによる指文字認識の検討、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第6回研究会、2018

橋本大空、松原正樹、白石優旗、張建偉、若月大輔、森嶋厚行、クラウドワーカ的能力とインクルージョン性を考慮したタスク割当て手法、第10回データ工学と情報マネジメントに関するフォーラム(DEIM 2018)、2018

畑伸佳、白石優旗、スマートフォンを用いた深層学習による警告音認識システムの検討、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第3回研究会、2017

設楽明寿、白石優旗、振動刺激を用いたクラウチングスタートにおける反応時間計測システムの開発、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第3回研究会、2017

若月大輔、張建偉、白石優旗、熊井克仁、森嶋厚行、クラウドソーシング型情報保障におけるタスク分割の基礎的検討、電子情報通信学会2017年総合大会、2017

設楽明寿、白石優旗、聴覚障害者陸上競技に適した振動刺激スタートシステムの提案、情報処理学会アクセシビリティ研究会(IPSJ SIG AAC)第2回研究会、2016

〔その他〕

ホームページ等

<https://www.shiraishi-lab.com/>

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。