

## 深層学習を用いた警告音認識による危険信号通知システムの開発

白石優旗

筑波技術大学 産業技術学部 産業情報学科

キーワード：情報保障, 警告音認識, 深層学習

### 1. はじめに

難聴高齢者を含む聴覚障害者が社会に進出し、安全・安心かつ質の高い生活（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。

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。

本研究では、識別対象には、日常生活における安全性の観点から、救急車のサイレン、並びに、歩行者や自転車の交通事故防止のためのクラクションやベルを選定し、識別手法に深層学習を採用することで、ノイズな環境音の中においても、対象物の移動や音響環境の変化による音質変化にロバストな高精度の識別を目指す（図1）。また、識別結果のユーザへの通知は、ウェアラブル端末への振動とスマートフォンへの画面表示を連動させて行うことで、確実かつ適切に情報伝達を行うことを目標とする。その際、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なシステム構想としている。

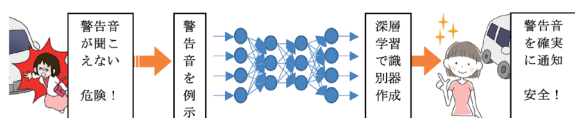


図1 提案システム

### 2. 関連研究

中西らは、音響モデルとする音声認識器 Julius を用いて環境音からその場の状況を判断の上、適切な情報を利用者に提示するモバイルアプリケーションの開発を行って

いる[1]。竹内らは、識別手法には IIR 型フィルタ組み合わせた閾値処理により、警告音を検知して使用者に知らせる歩行者向けのシステムを提案している[2]。松葉らは、LVQ により学習されたニューラルネットワークを用いて音源種の識別を行い、聴覚補助支援システムの開発している[3]。また、岩佐らは、パルスニューロンモデルによる識別も行っている[4]。

しかしながら、既存の研究では雑音環境下での識別精度が90%程度以下と不十分である。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る。本研究では、識別手法に深層学習を活用することで、これらの課題の解決を試みる。

一方で、Lane らは、DeepEar と呼ぶ、深層学習を用いたスマートフォン用の音響環境認識プラットフォームを開発している[5]。Ambient Scene Analysis や Stress Detection といった Audio Sensing Task では従来よりも高い識別精度を記録しているものの、警告音認識のような、安全性の観点から非常に高い精度を要求されるタスクの評価はなされていない。

### 3. 提案システム

本研究では、警告音識別時の通知方法としてウェアラブル端末による振動方式を採用する。例としては、スマートフォン（iPhone）とスマートウォッチ（Apple Watch）の組み合わせによる実現が考えられる。

提案システムの基本的な仕組みは以下の通りである。

1. スマートフォンにより環境音を収集
2. 警告音識別時にはスマートフォンにプッシュ通知
3. 警告音を確実に通知するためウェアラブル端末にも同時にプッシュ通知（振動等）

識別手法には深層学習を用いる。よって、多数の学習データの採取があらかじめ必要となる。その際、通知対象とする警告音に対して、スマートフォンを用いて多数の音デー

タを採取する。さらに、ノイズな環境音の中で、対象物の移動や音響環境の変化によって音質変化したデータも採取する。ここで、警告音の純音を採取するのではなく実環境下においてデータ採取する理由は、深層学習の汎化能力を最大限に活用するためである。

様々な環境下で採取した警告音データに対して、データ整理、並びに、データスクリーニングを行い、学習用データベースを作成した後、実際に学習を行う。

#### 4. 識別方法

警告音識別のためには

1. 連続的に環境音を集音し閾値以上の音量を検知
  2. 1. で検知された場合は一定時間の音データを記録
  3. 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となる。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ディープニューラルネットワーク (deep neural network, DNN) の入力とする。最後に、識別結果に統合処理を施すことでリアルタイム識別を可能とする。

#### 5. 予備実験

本稿では、救急車のサイレンと2種類の自転車のベルを対象とした識別実験を行い、深層学習の識別基本性能を確認する。音データの集取は実際にスマートフォン (iPhone 6s Plus) を用いて行った。なお、今回は閾値処理による音の切り出しは手動で行った。DNNの学習と評価はChainer (ver 1.6.1) を用いて行った。

今回の予備実験では、中間層の活性化関数をシグモイド関数とする従来型の3層 NN, 中間層の活性化関数を ReLU (ランプ関数) とする4層 DNN, 5層 DNN の3種類に対して学習及び評価を行った。誤差関数は、ソフトマックス交差エントロピーとし、学習アルゴリズムには Adam を用いた。

具体的には、3種の音源に対してそれぞれ9つの音データ (32bit, 44.1KHz) を採取し、6つを学習データ、残り3つを評価データとして評価した。その際、1024frame で STFT を行った。結果として、1350個のデータに対して、900個を学習に、残り450個を評価に用いた。なお、NNの入力層の素子数は513、中間層の素子数は128、出力層の素子数は3で統一した。1000エポックまで学習した結果、すべてのNNにおいて評価データの識別率100%を達成した。ただし、5層DNN, 4層DNN, 3層NNの順で、学習誤差並びに評価誤差が小さくなった。

#### 6. まとめと今後の課題

本研究は、環境音の中から緊急の回避や避難を必要とする警告音を識別し、聴覚障害者や聴力の低下した高齢者等に通知するシステムの開発を目的としたものであった。本稿では、救急車のサイレンと自転車のベルを識別対象とした予備実験により、提案システムの基本性能を確認した。今後は、識別対象の増加、通知システムの開発、聴覚障害者を対象としたユーザビリティ評価実験により、実用システムの実現を目指す。

#### 参考文献

- [1] 中西恭介, 津田貴彦, 西村竜一, 河原英紀, 入野俊夫. 環境音認識を応用した情報提供機能を有するモバイルアプリケーションの検討. 情報処理学会第75回全国大会, 2013.
- [2] 竹内浩一郎, 松本哲也, 竹内義則, 工藤博章, 大西昇. スマートフォンを用いた聴覚障害者のための警告音検知システム. 信学技報, WIT, Vol.113, No.481, pp.87-92, 2014.
- [3] 松葉亮太, クグレマウリシオ, 黒柳奨, 岩田彰. 聴覚補助支援システムの屋外使用モデルの開発. 信学技報, NC, Vol.111, No. 483, pp. 215-220, 2012.
- [4] 岩佐要, 藤角岳史, クグレマウリシオ, 黒柳奨, 岩田, 段野幹男, 宮治 正廣. 車載用安全運転支援装置のためのパルスニューロンモデルによる音源接近検出及び音源種類識別システム. 信学誌, D, Vol.91, No 4, pp.1130-1141, 2008.
- [5] N. D. Lane, P. Georgiev, L. Qendro. DeepEar: Robust Smartphone Audio Sensing in Unconstrained Acoustic Environments using Deep Learning. In Proc. of the UBIComp '15, Osaka, Japan, pp.283-294, 2015