

## 知的聴覚支援システムに関する研究

## Study on intelligent hearing assist system

梶本 阿斗

霜山 竜一

Ato Kajimoto

Ryuichi Shimoyama

日本大学大学院生産工学研究科

Graduate School of Industrial Technology, Nihon University

## 1. まえがき

聴覚障害者は外出時に側方や後方から接近する自動車やバイクに気づかず危険にさらされる場合がある。著者らが開発した、音源の位置を推定し障害者に呈示するウェアラブル聴覚支援システム<sup>1)</sup>は、風など他の環境音にも反応するため、屋外で動作が不安定になる。環境音をAIで認識させる試みの多くは室内でなされており、屋外で遠方の車両の走行音をAIで認識させた例はない。

本報告では、音を認識することで特定の音源のみに反応して聴覚支援を行わせるため、USB Acceleratorを用いたエッジAIで環境音の認識を試みた。音質に応じて危険か安全かを区別した。

学習済みAIプログラムをPCに実装し、従来の聴覚支援システムに接続することで音質に応じて障害者を支援する。システム全体の構成を図3に示す。エッジAIの出力信号は“0”(安全)か“1”(危険)で、“1”の場合のみ聴覚支援を行う。



Fig.1 Flowchart  
(AutoML Vision API)

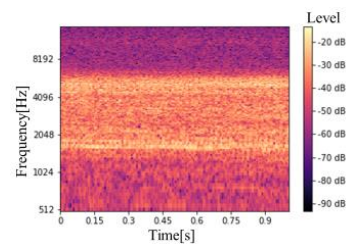


Fig.2 Frequency spectrogram (Motorcycle)

## 2. GCP と USB Accelerator について

Google Cloud Platform(GCP)の AutoML Vision API を使用して環境音を画像として認識させた。AutoML Vision API はプログラムの起動に時間を要するうえ、常時接続すると利用料金が高額になるためリアルタイム処理には適さない。そこでGCPで開発したAPIをUSB Accelerator(Coral)付きのPCで動作させ、ローカル環境でAIによる音の認識を試みた。

## 3. 処理の流れ

AutoML Vision APIを用いて環境音の認識を行うための処理の流れを図1に示す。インターネットや実際に録音して得た音響波形のwaveファイルを周波数スペクトログラムに変換した(図2)。測定時間を1[s]、測定周波数帯域を500[Hz]~15[kHz]に設定した。音のレベル差が画像に及ぼす影響を抑制するため音のレベルを規格化して配色した。また、遠くまで伝搬しやすい低周波数の音の成分を強調するために周波数軸を対数表示した。この周波数スペクトログラムを画像とみなして9ラベルに分類したものをAutoML Vision APIに学習させた。一方、マイクロホンから入力した音をwavファイル化した後に画像化して、学習済みAIに認識させた。各ラベルの認識率を求め最も認識率の高いラベル名を出力した。

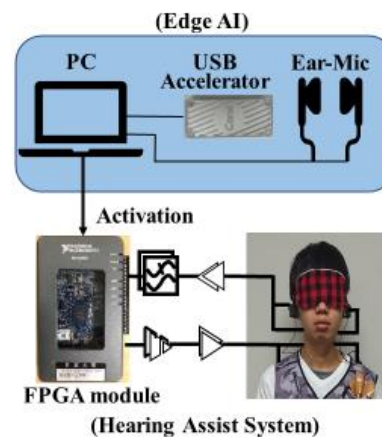


Fig.3 Configuration  
(Hearing assist system with Edge AI)

Study on intelligent hearing assist system

Ato Kajimoto and Ryuichi Shimoyama

学習用の画像を図4に示す。学習させた音のラベルは自動二輪車、バス、普通乗用車、自然音、生活音、動物音、人の声、風の音、環境ノイズの9種類である。画像数は自動二輪車が300枚、バス200枚、普通乗用車200枚、自然音100枚、生活音100枚、動物の音100枚、人の声100枚、風の音100枚、環境ノイズ100枚の計1300枚である。音のラベルを危険か安全かで分類した(表1)。自動二輪車、バス、普通乗用車を危険と分類し、動物の音、自然音、生活音、人の声、風の音、環境ノイズを安全と分類した。

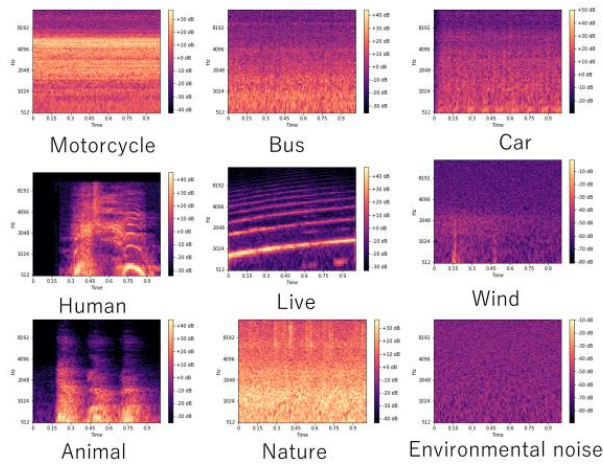


Fig.4 Training Images (9Labels)

Table.1 Image Classification (Danger or Safe)

Motorcycle	Danger
Bus	Danger
Car	Danger
Animal	Safe
Human	Safe
Live	Safe
Nature	Safe
Wind	Safe
Environmental noise	Safe

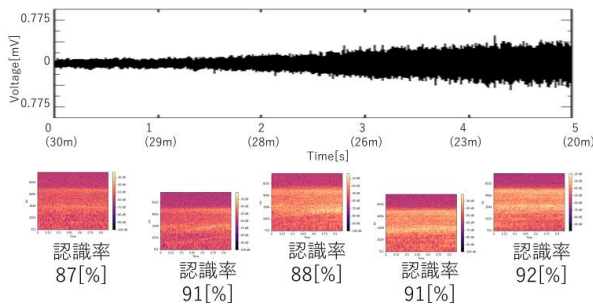


Fig.5 Sound recognition of traveling motorcycle(GCP)

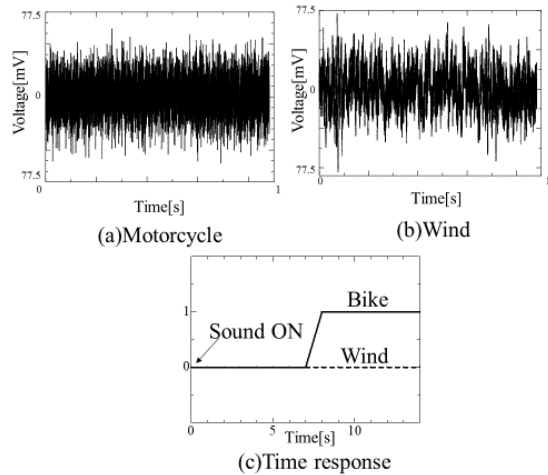


Fig.6 Time response of output signal of Edge AI

#### 4.実験結果

屋外でマイクロホンから30[m]離れた位置から10[m]手前まで自動二輪車を走行させた。録音した音の波形を画像化して、Cloud-base AIで認識させた。走行時の自動二輪車のエンジン音の時間波形と1[s]毎の認識率を図5に示す。バイクが接近するにつれて音の波形の振幅が大きくなるのがわかる。周波数スペクトログラムは自動二輪車が接近してもパターンに変化がなく、認識率も全ての区間で80[%]以上を示した。

スピーカから各々自動二輪車と風の音を発生させた場合に、エッジAIから出力される信号の時間応答を図6に示す。同図(a)は自動二輪車のエンジン音、同図(b)は風の音の波形である。同図(c)はエッジAIの出力信号の時間経過である。自動二輪車の音では7秒後に出力が0から1に変化した。風の音では信号は変化しなかった。エッジAIによって表1に定義される危険な音を認識/区別できた。

#### 5.おわりに

危険な音のみに反応して聴覚支援を行う目的でエッジAIによって、予め設定した音(危険/安全)を認識/区別できることを示した。今後は7秒程度要する応答時間を短縮し、エッジAIの出力をFPGAモジュールに入力し、聴覚支援システムを動作させる予定である。

#### 参考文献

- 1) 関淳ら：両耳聴型聴覚支援システムにおける遠方の音源の方向呈示について研究，電子情報通信学会2020年総合大会,2020-3,H-4-7,p.232