

音による救助支援システム「Acous」

venv

田中優陽

溝辺陸

白川息吹



目次

- 01 事業コンセプト
- 02 プロトタイプ
- 03 ディープレアニングについて
- 04 ハードウェアについて
- 05 ビジネスプラン

01

事業コンセプト

解決したい課題

災害現場では、被災者の声や音が届かず、発見が難しい

具体的には以下のような課題があります。

救助隊

視界が遮られた建物内や土砂崩れ現場では、目視や声に頼った搜索に限界があり、被災者の位置を特定することが難しい。

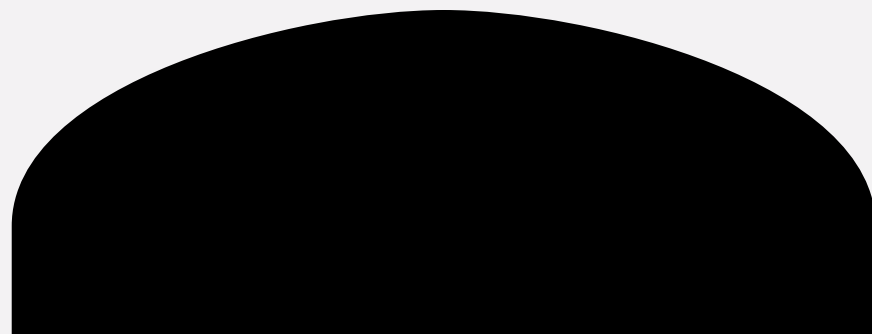
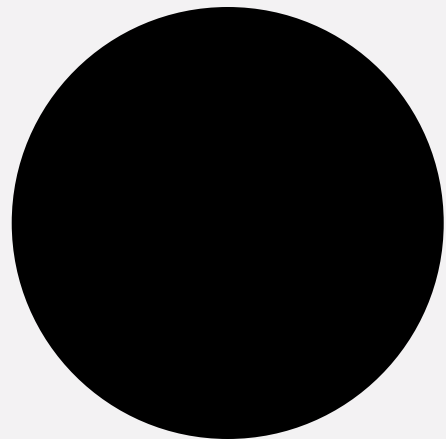
被災者

建物内や瓦礫の下に取り残されても、声が出せない、または助けを求める音が届かず、自身の存在を伝えられない。

01

事業コンセプト

作品に取り組んだ背景・理由



リーダー 田中

ニュースで大きな地震や火災の映像を見るたびに、もし自分がその場に取り残されたらと想像しました。
声を出せず建物に閉じ込められ、誰にも気づいてもらえない状況は、大きな恐怖を伴います。
こうした背景から、将来起こりうる災害に備え、一人でも多くの命を救うためのプロダクトを開発したいと考えました。

01

事業コンセプト

現場の声

現場での救助活動について詳しく知るため、寝屋川消防署を訪問し、救助隊の方からお話を伺いました。その結果、以下のようなことがわかりました。

倒壊現場では…

救助隊が声掛けを行い、
反応の有無を確認

その後



人がいる可能性がある場合には、
地中音響探知機を用いて被災者の発する音を捉え、
位置の特定を行う

しかし、地中音響探知機には以下のような課題があります。

- 被災者の反応が前提であり、反応がなければ使用されない
- 拾った音を大きさとしてしか扱えず、その音が人によるものかどうかを判断できない
- 騒音が多い災害現場では、ノイズの影響を受けやすい
- 被災者の位置を特定するために、装置の設置場所を何度も変える必要がある
- 装置が重いため持ち運びが救助隊の負担になる（装置は約30kgある）
- 装置の価格が高く、十分な台数を導入することが難しい（装置の価格は約300万円）

01

事業コンセプト 現場の声

また、次のようなこともわかりました。

災害現場での救助活動の流れ

害状況を把握するために
先遣隊が現場へ派遣される

その後



状況確認が終わってから救助隊が派遣され、
被災者の捜索・救助が行われる

そのため・・・

災害発生から救助隊の活動開始までに時間がかかってしまう。

さらに、救助隊が到着してから地中音響探知機などの機器を使って捜索を始めるため、被災者の発見までにさらに時間がかかる。

02 プロトタイプ プロダクトの概要

そこで私たちが開発したのが、

倒壊建物内部の音をAIで解析し、人の存在を検知する救助支援システム

Acous (アコース)



02 プロトタイプ

プロダクトの特長

「Acous」には、次のような特長があります。

倒壊した建物内部の音をAIが解析し、被災者がいるかどうかを自動で判断

救助隊の方がその場に張り付いて音の反応を確認し続ける必要がなく、設置するだけで使用できる

地中音響探知機と比べて導入コストが低い

より多くの台数を現場に配備でき、同時に広い範囲を調査できる

装置が小型・軽量

救助隊の方の身体的負担を減らせるだけでなく、一度に複数台を持ち運ぶことができる

倒壊した建物の瓦礫の中に設置できる

瓦礫の中に入れて設置することで要救助者との距離を縮め、微かな音でも発見しやすくなる
また、瓦礫の上でも安定して設置できる構造としている

02 プロトタイプ プロダクトの構成

「Acous」は以下のように構成されています。



02 プロトタイプ

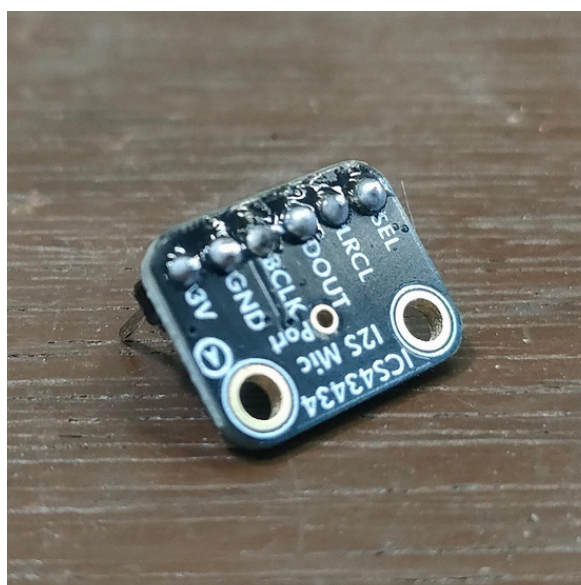
プロダクトのワークフロー

音の取得から被災者の検知・表示までの流れを示したシステム構成図です。

收音部

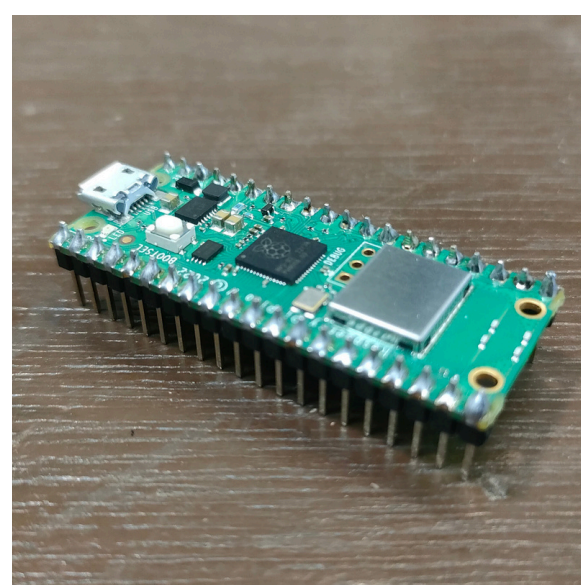
処理部

マイク



音の収集

マイコン



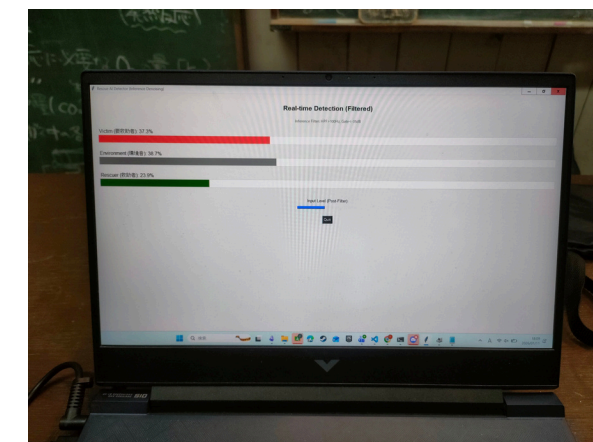
音データの取得
整形・転送

ノートPC



AIによる解析

確認画面



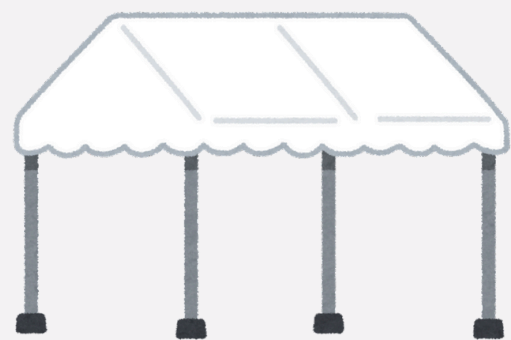
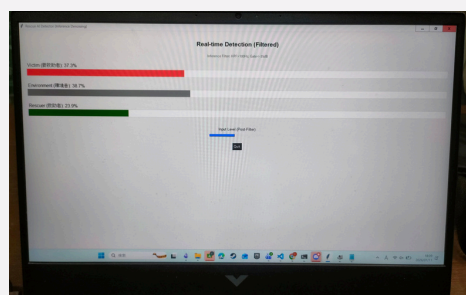
解析結果の表示

02 プロトタイプ

プロダクトの使い方

災害現場では、次の方法で「Acous」を設置・運用します。

災害対策本部



現場



1. 初動調査と同時に設置

災害発生後、被害状況を確認するために最初に現場へ入る先遣隊が、初動調査とあわせて「Acous」を設置します。

2. 置くだけで自動解析

Acousは、收音部を瓦礫の隙間から建物内部に入れ、処理部を地上に設置するだけで使用できます。

3. 検知結果を本部へ共有

設置後は、AIが音を自動で解析し、人の存在を検知した場合、その情報が災害対策本部に送信され、救助判断に活用されます。

02 プロトタイプ

今後の計画

現在はまだ未実装ですが、今後取り組む予定の内容です。

処理部の開発と災害対策本部への情報送信

現在はAIによる解析をノートPCで行っていますが、軽量化や携帯性の向上、コスト削減を目的として、専用の処理部を新たに製作します。また、現状では人の存在を検知した結果を処理部(パソコン)上で確認していますが、今後は解析結果を災害対策本部へ送信し、遠隔から確認できるようにします。

收音部の複数化と連携

現在はAcous1セットにつき收音部が1つですが、今後は收音部を複数設置し、それぞれから得られる音を組み合わせることで、より多くの被災者を発見し、位置を特定できるようにします。

マイクのハイブリッド化

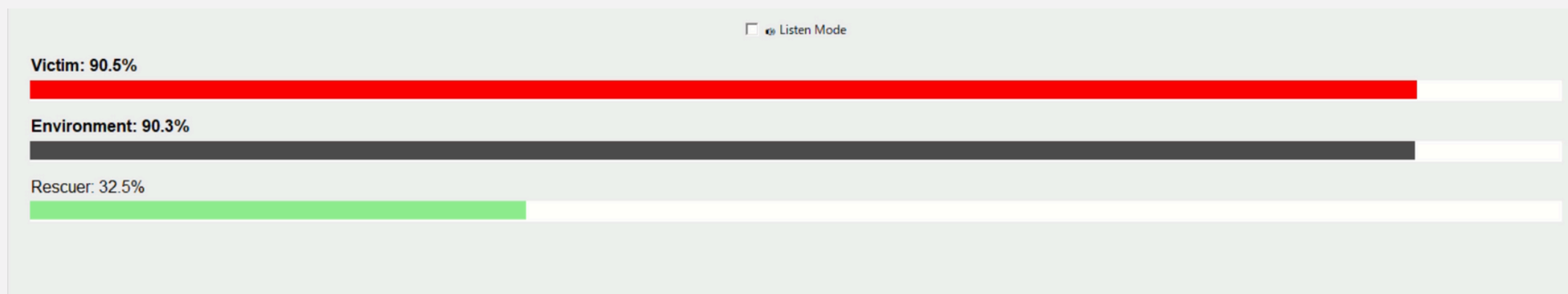
現在は收音部1つにつきMEMSマイクを1つ搭載していますが、今後はそれに加えて振動センサを併用することで、より幅広いアプローチで要救助者の発見を目指します。

03 ディープラーニングについて

ディープラーニングの利用

私たちは、環境雑音に埋もれた瓦礫下の要救助者によるSOS(打撃音)をリアルタイムで検出するため、ディープラーニングモデルを構築しました。

このモデルは、收音デバイスからPCに入力された音声信号に対し、システムは「Victim(要救助者)」「Environment(環境音)」「Rescuer(救助者)」と定義された3クラスそれぞれの存在確率を算出し、出力します。

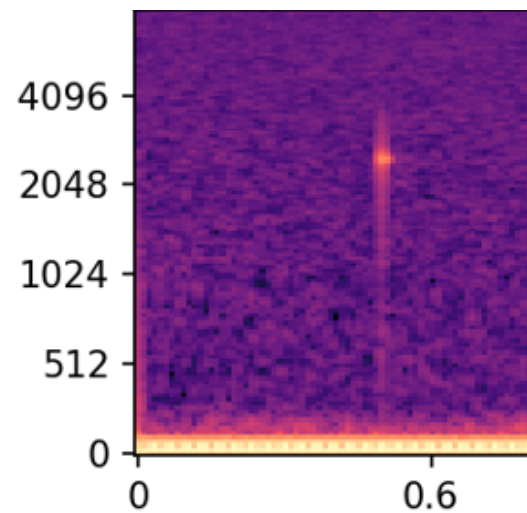


03

ディープラーニングについて AI解析の処理フロー（パイプライン）

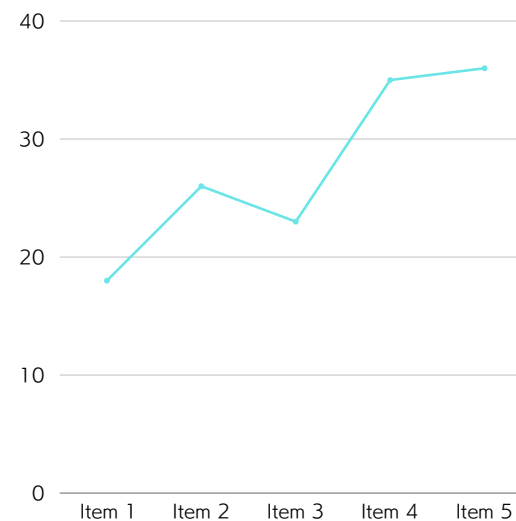
🔊 音の収集

前処理



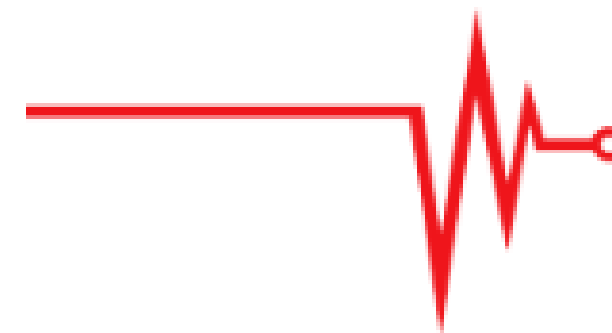
音を画像化

CNN



特徴抽出

GRU



時系列解析

確率を出す

ノック音+環境音

Victim: 90%

Environment: 99%

Rescuer: 20%

解析結果の表示

03 ディープラーニングについて

使用したモデルと選定理由

AIが解析しやすい形式に音を画像化したメルスペクトログラムをモデルの入力としました。

使用モデル



1. 空間的特徴抽出

3層の畳み込み層(Conv2d)を採用し、メルスペクトログラムから音響イベントの空間的特徴を抽出しています。具体的には、ノッキング音がスペクトル上で形成する「垂直方向の鋭い線」を画像として捉えることで、瓦礫による周波数減衰の影響を受けやすい「音色」ではなく、堅牢な「信号形状」に基づく識別を行っています。

2. 時間的文脈学習

自然崩落による単発の衝撃音と、人間による意図的なSOS(規則的なノック)を区別するため、双方向GRUで音の周期性・リズムを学習させています。また、救助隊が発する「大丈夫ですか」などの文脈を理解するためにも使用しています。

3. クラス分類

抽出された時空間特徴を全結合層(Fully Connected Layer)に入力し、「Victim(要救助者)」「Environment(環境音)」「Rescuer(救助者)」の3クラスに対する存在確率を出力しています。

CRNNという比較的軽量なアーキテクチャを用いることで、高速な推論(約3.0ms)を可能にしており、リアルタイム音声分類においても、遅延がほぼなしの状態で作動させることができます。

03 ディープラーニングについて

学習データ・テストデータの収集方法・ラベリング

学習・検証に使用したデータ(計205個)は以下の方法で収集しました。

自分たちで録音して収集

- 様々な物質のノック音
- 暖房、風の音などの環境音
- 「誰かいませんかー」などの救助隊が言いそうな言葉

+

インターネットで収集

- 様々な物質のノック音
- 雨音、ドローンの音、重機などの環境音
- 実際の救助隊の訓練映像に含まれていた音声

収集したデータセットは以下のようにクラス分けを行いました。

要救助者 (Victim): 壁や床を叩くノック音(SOS打撃音)

環境音 (Environment): 重機の駆動音、防災訓練時のサイレンや喧騒、風雨の音

救助者 (Rescuer): 呼びかけの声

データは以下のような形式でマルチラベリングを行いました。

filename(ファイル名),label(名前でのラベル付け),environment,rescuer,victim(数字でのラベル付け)

car_res01.wav,rescuer,0,1,0

03 ディープラーニングについて

学習戦略について

• 学習データ

収集した205個のデータは、学習データ(80%, 164個)と検証データ(20%, 41個)に分割しています。

データのクラスごとの内訳はVictim:88個, Environment:184個, Rescuer:66個です(マルチラベルのため足しても205個にはなりません) 学習データにはノイズ付与、ピッチシフトなどのData Augmentationを適用し、実質的な学習データ数を確保して汎化性能を高めています。

• マルチラベル分類

災害現場の音環境は「排他的」でないという問題があります。「激しい雨音(Environment)」の中で「救助を求めるノック(Victim)」が鳴るなどの複数の事象が同時発生します。

そこで、各クラスの独立性を担保するため、出力層の活性化関数にシグモイド関数を採用し、損失関数にはBCEWithLogitsLossを使用しました。

これにより、softmaxでは埋もれてしまうような、高ノイズ下での微弱なSOS信号の検知、「Environment 0.95」かつ「Victim 0.85」などの重複した音声の検知が可能になりました。

• データ不均衡への対策

収集したデータにおいて、自分たちで音声を収録する際に、必然的に環境音が入力されてしまうため、環境音クラスの発生頻度が高くなってしまふことが課題でした。これに対し、出現頻度が比較的少ないクラス(victim, rescuer)の学習重みを高く設定する「逆頻度重みづけ」を適用した BCE Lossを採用しています。これにより、すべてのクラスを公平に学習させています。

03 ディープラーニングについて

性能評価

• Marco-averaged F1-Scoreの導入

本タスクはマルチラベル分類であり、特に環境音クラスに比べて、他の2クラスの発生頻度が少ない不均衡データとなっています。そのため、評価指標として単純な正解率ではなく、「正解である事象をどれだけ正確に捉えられたか」を重視するF1スコアを採用し、さらにクラス間のデータ数差に左右されず、全クラスを公平に評価するために、Marco-averageを用いています。

また、人命救助においては、誤検知(Precision低下)のコストよりも、救助の遅れや見逃し(Recall低下)が最も致命的となります。そこで、総合指標としてのF1スコアを維持しつつ、見逃しリスクを最小化する運用設定として、各クラスに対し独立した比較的小さな閾値判定(0.4)を行っています。

この戦略により、現状のモデルは誤検知の課題を残しつつも、要救助者の見逃しという最大のリスクを極限まで低減することに成功しています。

クラス (Class)	Precision (誤検知の少なさ)	Recall (見逃しの少なさ)	F1 Score (調和平均)	備考
victim (要救助者)	0.64	0.941	0.762	Recall94.1%を達成
environment (環境)	0.919	1	0.958	ほぼ完璧に分類
rescuer (救助者)	0.583	0.875	0.7	誤検知が多い

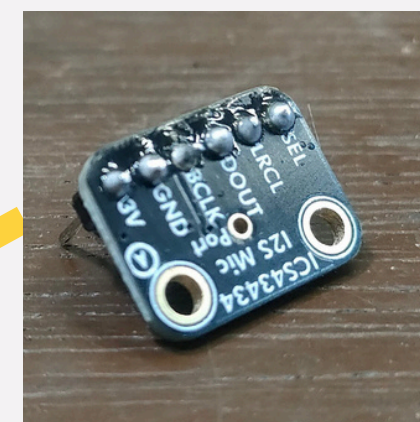
表 各クラスに対するPrecision、Recall、F1 Score

04 ハードウェアについて

ハードウェアの構成

ハードウェアとして、主に收音部を製作しました。その構成は以下のとおりです。收音部を処理部であるノートPCと接続することで、収集した音をディープラーニングで解析できます。

マイコン



マイク

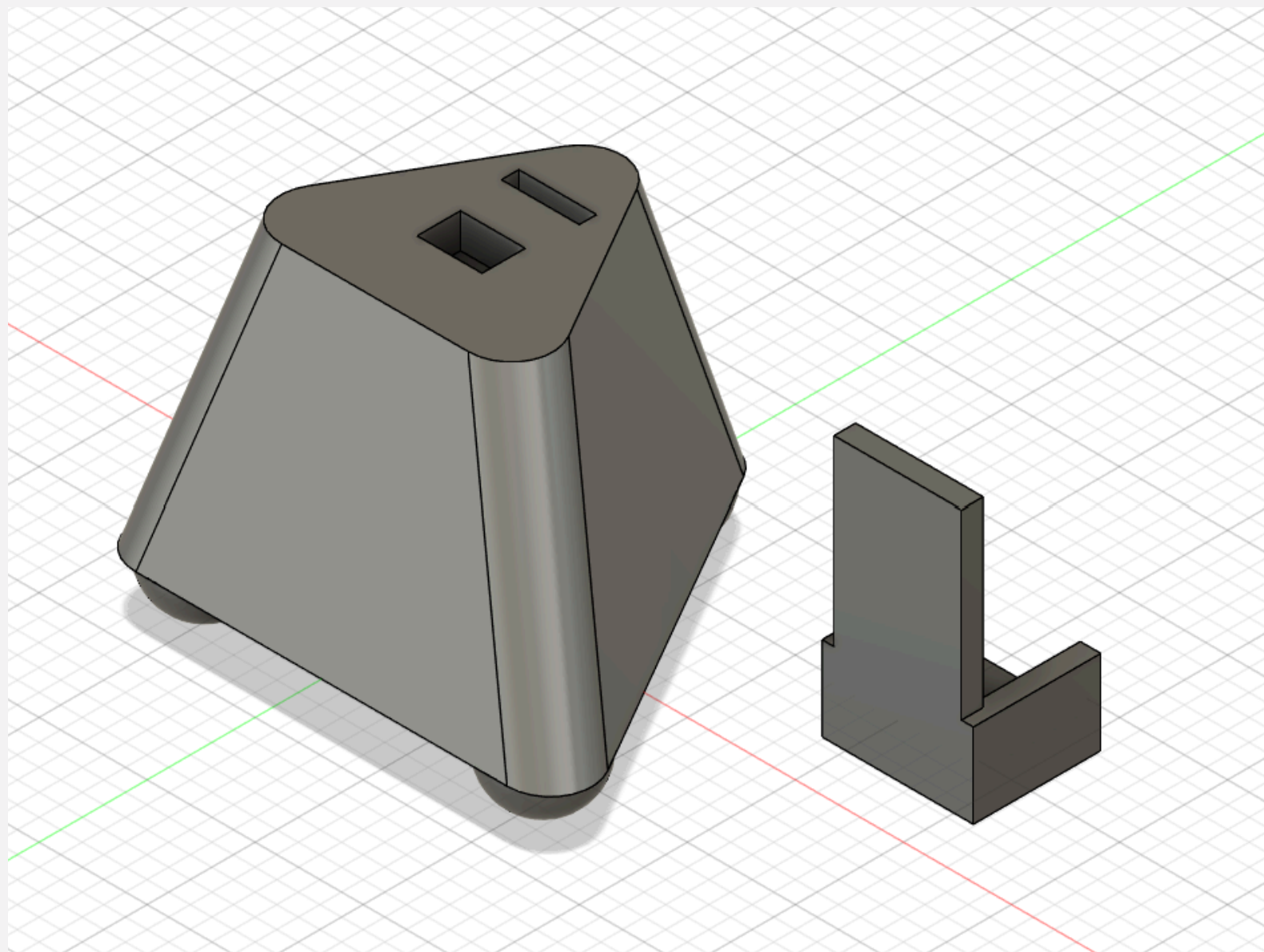


土台

04 ハードウェアについて

製作したハードウェア

收音部の土台は、3DCADを用いて設計し、3Dプリンタで製作しました。



3DCADで設計した收音部土台



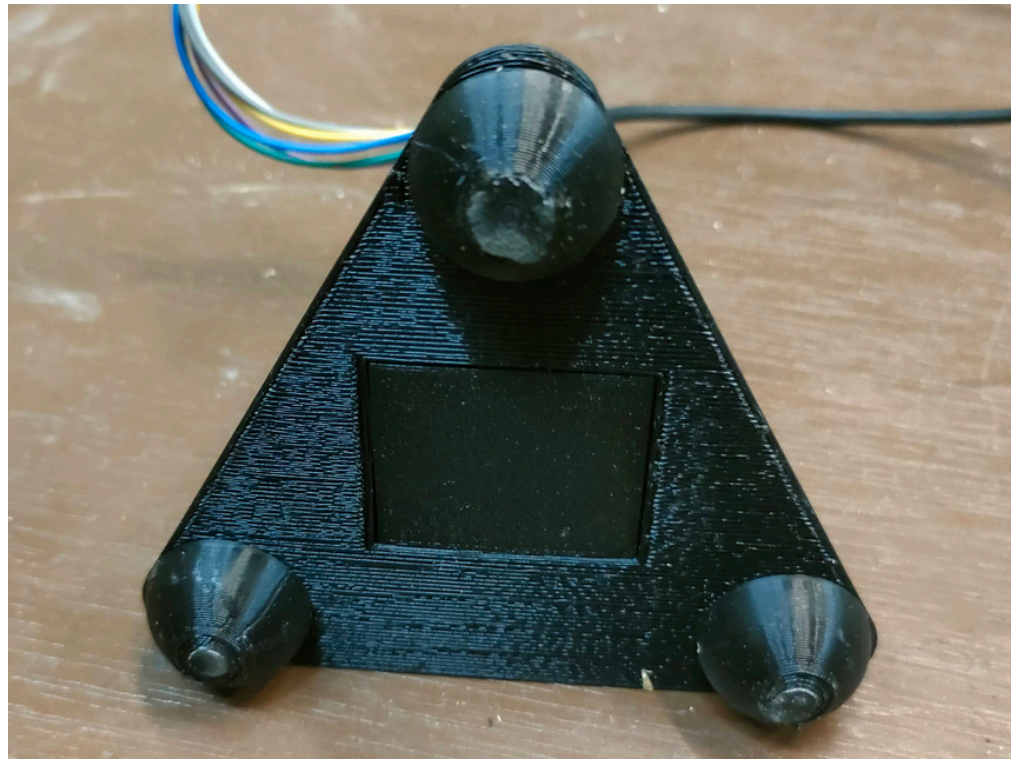
3Dプリンタで製作した收音部土台

04 ハードウェアについて

工夫した点

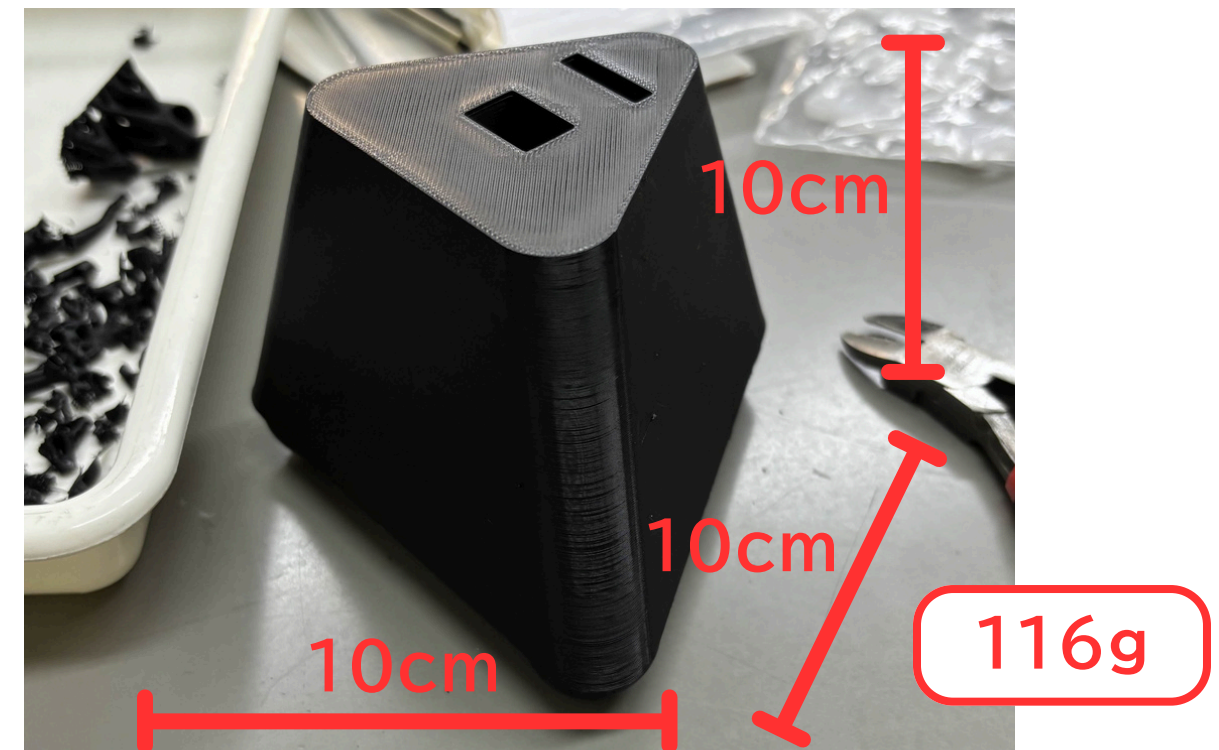
收音部には、災害現場での使用を想定した2つの工夫があります。

安定して設置できる構造



瓦礫の上など地面が不安定な場所でも設置できるように、3点で支える三本足構造を採用しています。

小型・軽量化



土台の大きさは、内部に必要な部品を収めつつ、安定性を確保できる最小限のサイズとしています。

05

ビジネスプラン

ターゲット市場と収益モデル

以下に、ターゲット顧客・価格設定・想定される市場規模を示します。

ターゲット顧客

地方自治体

提供価値

災害時に住民をより早く発見できる可能性を高める

価格設定

収音部1台 + AI解析プログラムを1セットとし、

1セットあたり3万円

(処理部ハードは販売しない)

コスト

収音部の制作費用

1セットあたり約3千円

想定される市場規模

$1,700$ (日本の自治体数) \times 3 万円 \times 50 台 = 25.5 億円

05

ビジネスプラン

競合サービスと市場の現状

救助支援／災害現場向けの既存プロダクトの例を示します。

地中音響探知機

倒壊建物内の音・振動から
生存者の存在を探す従来の救助機器

総合救助キット

地中音響・振動センサーと
探索カメラを組み合わせた救助装備

UWBレーダー救助装置

微細な動きを捉える
レーダータイプの救助支援機器

RECCO システム

遭難者追跡用の反射／検知システム
(雪崩等)

05

ビジネスプラン

競合優位性とAcousの強み

一方で、Acousには次のような強みがあります。

ディープラーニングによる高い検知性能

Acousはディープラーニングを用いて音を解析するため、高性能なマイクなどの装備に頼らず、必要最小限の構成でも被災者の発見につなげることができます。その結果、コストを抑えつつ、小型・軽量化を実現できます。

また、雑音の多い災害現場でも音の特徴を判別できるため、周囲の環境音に影響されやすく、判断が難しくなる従来製品よりも安定した検知が可能です。

被災者が事前に準備する必要がない

Acousは、被災者が特別な機器や準備をしていなくても、発した音や物音から人の存在を検知できます。

一方で、既存の製品には反射器や専用機器を被災者自身が常に持ち歩いていることを前提としたものもあります。

Acousはそのような前提に依らないため、準備ができていない状況でも検索が可能です。

「Acous」を通して、

災害が起きたとしても、誰一人として命を落とさない社会を実現します