

# 屋外環境下でのドローンのローターノイズによる地表材質推定に向けた手法の検討およびマイクロホンアレイ用風防の設計

## Method for Ground Surface Material Classification using Drone Rotor Noise in Outdoor Environment and Development of a Windshield for Microphone Array

矢野 翼<sup>1\*</sup> Benjamin Yen<sup>1</sup> 中臺 一博<sup>1</sup>  
Tsubasa Yano<sup>1</sup> Benjamin Yen<sup>1</sup> Kazuhiro Nakadai

<sup>1</sup> 東京科学大学

<sup>1</sup> Institute of Science Tokyo

**Abstract:** 本稿では、ドローンから発生するローターノイズを用いて、ドローン直下の地表面の材質を推定する手法について扱う。この手法により、人が容易に近づけない場所での被害状況を評価できるようになることが期待される。風などの外来ノイズが存在する実環境において、アスファルト、土砂、水などの複数の地表材質上でドローンノイズを収録し、深層学習ベースのモデルを用いて地表材質の分類を実施した。また、外来ノイズの影響を抑えるべく、マイクロホンアレイ用の風防の開発も行った。その結果、分類手法と風防双方の有効性が確認できた。

## 1 はじめに

近年、豪雨による土砂災害や地震などの自然災害が頻発しており、たびたび大きな被害をもたらしている。日本では、2019年の台風19号で東日本を中心に20都県にわたって950件を超える土砂災害が発生し[1]、2020年7月に熊本県を襲った豪雨では、河川の氾濫や土砂崩れに伴う国道の通行止めが複数件発生した。さらに、直近の2024年1月に発生した能登半島地震では、県が管理する道路だけでも90か所以上で通行止めとなった。また、半島という地理的要因も相まって、孤立集落の発生状況は、震災発生から3日経ってなお詳細不明な地区が存在するなど、被災状況の把握に時間を要した[2]。異常気象などにより引き起こされる災害は日本のみならず世界中で増加傾向にあり[3]、それらへの対策と対処が求められる。

このような問題を解決する手段の一つが、機動性の高いドローンを利用することである。ドローンを用いれば、人が直接立ち入ることが難しい場所についても迅速かつ低コストでの調査が可能になることが期待できる。

そこで、本稿では、ドローンに搭載したマイクロホンで収録する音響信号には、ドローンのロータから直接届く雑音に加え、一度地表に反射してから届く反射

音も含まれていることに着目し、この反射音を利用した地表材質推定問題を扱う。また、屋外で地表材質を推定するためのデータセットが存在しないことから、このようなデータセットの作成も併せて実施する。また、屋外では強風による外来ノイズの影響が懸念されることから、本稿で使用するマイクロホンアレイ用の特注の風防を新しく提案し、様々な素材を用いた風防との比較を通じて、その性能についても検証を行う。

## 2 関連研究

ドローンにセンサを搭載し、被災状況を把握する研究の多くは、センサとしてカメラやマイクロホンを利用している[4]。カメラは、状況把握に有効なセンサであり、実際に被災地でも利用されているが、夜間など視界が悪い条件では使用が制限される。マイクロホンは、こうしたカメラの問題を解決できることから「ドローン聴覚」として研究が行われている[5, 6, 7, 8, 9, 10]。しかし、ドローン聴覚では、多くが被災者由来の声を検出するために、ドローンのローターノイズを抑圧するアプローチをとっている。つまり、ドローンのローターノイズを如何に抑圧するかという文脈で研究が行われている。一方、本稿では、ドローンのローターノイズを抑圧する対象としてではなく、地表面の材質を知るための有用な情報として積極的に活用するという点で一線を画しており、まったく異なるアプローチによ

\*連絡先：東京科学大学  
東京都目黒区大岡山 2-12-1  
E-mail: yano@ra.sc.e.titech.ac.jp



図 1: 本稿の応用例のイメージ

るドローン聴覚研究であるといえる。このような地表面の材質推定ができれば、例えば図 1 のように、地表面物質や特徴の変化を特定することで、地滑りや道路崩壊を検出できる可能性が広がる。これは被災状況を把握するのに大いに役立つことが期待される。このような研究は、われわれが知る限り、われわれの研究を除き、ほとんど行われていない。また、われわれの研究 [11] にしても、地表材質を変えながら収録した雑音を用いて地表面の材質推定を行っているものの、データ収録が無響室である、風切り音などのドローンのローターノイズ以外の雑音を考慮していない、データ量も 1 分間程度の極めて少量のデータであるなど、実環境で有効であることを示すには至っていない。

音響信号から材質を推定する研究としては、橋梁構造物の打音検査があげられる。一般に、こうした検査では、熟練者が橋梁に打撃を与えたときに発生する音から、橋梁構造物の特性や劣化状態を推定していたが、近年では、熟練者の技術を機械学習的手法で学習し、同等の性能を得ようとする研究が盛んに行われている [12]。中には、打音検査機能を取り付けた付着走行型ドローンを用いた事例も報告されている [13]。しかし、これらの研究は、コンクリートを対象にした研究であり、様々な材質を反射音から識別するといった観点での研究は行われていない。また、打音検査機能付きドローンにしても、ドローンのローターノイズは不要な背景雑音として扱われており、本稿のように積極的にドローン雑音を利用するという視点は欠けている。

### 3 データセットの構築

#### 3.1 データセット設計指針

録音されるドローンのローターノイズに影響を与える要素は、地表の材質のほか、ドローンの高度、各モーターの回転速度、地表の形状など多岐にわたる。本稿では、屋外環境への応用の第一歩として、以下の指針の下でデータセットの作成を行う。

1. 地表の材質とドローンの高度以外のパラメータを固定する。
2. 地表の材質には、実際の災害状況を想定したものを選択する。

表 1: データセット作成条件

使用機体	DJI Matrice 210
アレイチャンネル数	16 (球形)
サンプリング周波数	16 kHz

3. ドローンノイズの反射音が減衰することで材質推定が困難になることが予想されるため、ドローンの高度には、実際の災害救助の場面と同等の高さと、それよりも低く、かつ人間の頭上まで余裕のある高さの 2 種類を設定する。
4. データ量を確保するため、各材質に対する収録時間については、ドローンのバッテリーの範囲内で可能な限り長時間収録する。

#### 3.2 データセット作成手法

図 2 に示すように、ドローンの前方向に突き出す形で、16 ch の球形マイクロホンアレイを、後方部には、録音保存用の Raspberry Pi をドローンの重心を考慮して、設置位置を調整して固定する。この状態でプロペラを回転させ、地表の材質を都度変えながら、取り付け付けたマイクロホンアレイの全てのチャンネルで録音を行う。

地表の材質には、土砂、アスファルト、水を選択した。水については深さ 0.7m (以後「浅い水」と深さ 5m (以後「深い水」) の 2 種類を用意し、全体で 4 種類の材質を選択した。いずれの地表面も、傾斜や大きな凹凸がなく、ほぼ水平である。各地表面の様子を図 3 に示す。ドローンの高度は、先に示した指針に従い、5m と 10m の 2 種類を選択した。

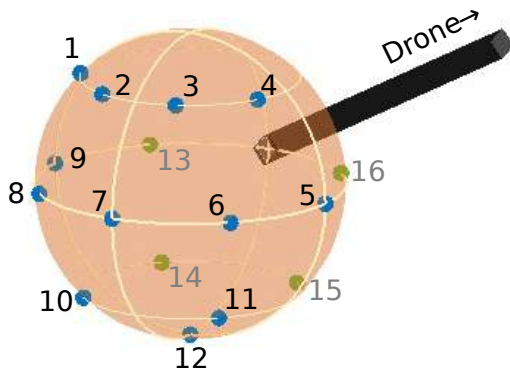
収録手順としては、まず各材質の上空 5m の位置でホバリングしたまま 8 分間録音を行い、録音終了後そのまま高度 10m の位置まで上昇し、ホバリングしたまま 8 分間録音を行った。安全のため、ドローンの低電池残量を知らせるアラームが途中で鳴った場合は、収録作業を打ち切り、速やかにドローンを着陸させるものとした。実際、浅い水の高度 10m では、ドローンの低電池残量を知らせるアラームが途中で鳴ったため、6 分 20 秒間で録音を中断した。原因としては、強風の影響を受け、想定よりもドローンのバッテリー消費量が大きかったためと考えられる。最終的に約 1h のデータを収録することができた。その他の条件については、表 1 に示す通りである。

#### 3.3 収録音の分析

録音したドローンノイズに異常や不具合がないかを確認するべく、聴取のほか、スペクトログラムと波形の



(a) 機器取り付け後のドローン



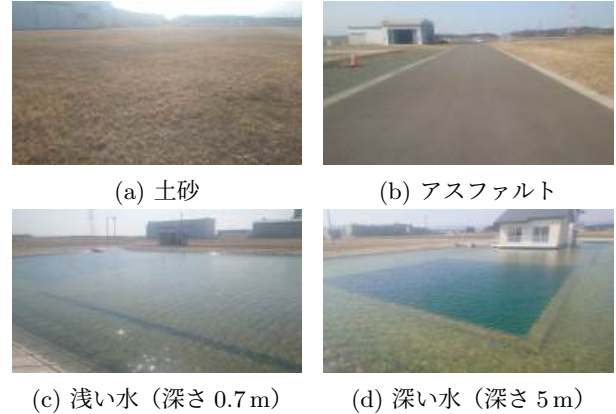
(b) アレイチャンネル配置

図 2: 収録に用いるドローンとマイクアレイ

確認をしたところ、以下2点の不具合が明らかになった。

- 録音データに頻繁に音割れが発生
- 録音データの値が0もしくは0に極めて近くなる区間が発生

図 4 は浅い水の上空で収録した第 14 チャンネル（後方側に位置）の録音のスペクトログラムであり、明るい色ほど強度が強いことを表している。低周波帯から高周波帯に渡って、インパルス的な信号が度々記録されていることがわかる。このようなインパルス的な雑音による音割れは、多くのチャンネルで共通の時刻に発生している。このことから、音割れの原因は、収録時の強風であると考えられる。このような音割れが発生してしまうと、反射音の信号を適切に拾えなくなるため、推定性能が劣化することが考えられる。一方、これとは対照的に録音データの値が 0 になる現象を図 5 に示す。これは土砂の上空で収録した第 10 チャンネルと第 11 チャンネルの録音のスペクトログラムであり、特に第 11 チャンネルでは全体の 9 割近くで値が 0 になっている。この現象が発生する時間長や頻度はチャンネルごとに異なっている。マイクロホンのメーカーにも相談したところ、強風により、利用している MEMS マイクロホンには、強い風が直接あたると振動板が内壁



(c) 浅い水（深さ 0.7m） (d) 深い水（深さ 5m）

図 3: 各地表面の様子

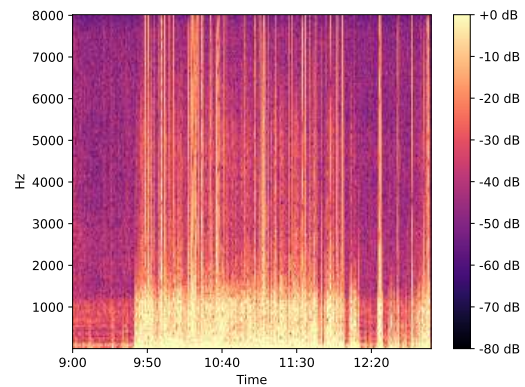


図 4: 音割れ発生状況

に貼りついた状態になり、感度が極めて小さくなってしまふことがあることがわかった。この現象も、強風によるものであることが考えられる。実際、収録時は風が強く、手元の風速計では平均 4-7m/s 程度、最大で 17 m/s を観測していた。以上の不具合から、本データセットには地表材質推定に利用できない部分が多いため、使用可能な部分を取り出すものとした。具体的な方法に関しては、5.1.1 節にて説明する。

## 4 マイクロホンアレイ用風防の設計

前節でも言及した通り、屋外環境下で地表材質推定を行うには、強風をはじめとする外来ノイズへの対処が必要となり、マイクロホンアレイへの風防の取り付けはその一つの手段である。多孔質構造の材料は吸音性に優れていることが知られており、特にポリウレタンを用いたものは、土木建築や交通輸送などの業界で広く利用されている [14]。そこでわれわれは、図 6 に示す、多孔質ポリウレタンフォームからなる風防を提案する。この風防は 2 つのパーツを組み合わせること

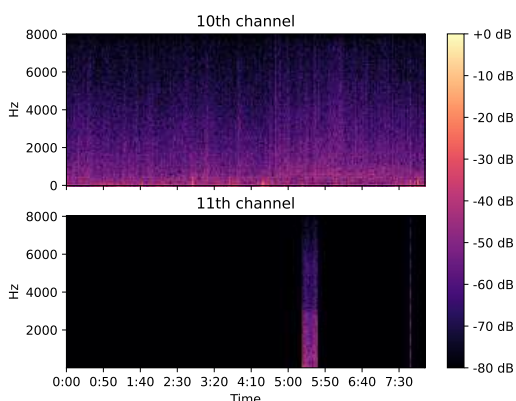


図 5: データ値 0 区間

でマイクロホンアレイ全体を覆うことができるようになってきている。また、この素材は水と合わせることで体積が 100 倍に膨れ上がる性質を有しており、風防内部に多くの気泡を含んでいる。そのため、全体で 21g と非常に軽量であり、荷重制限のあるドローンに対しても利用しやすいものとなっているほか、高い吸音性が期待できる。本稿では、実際に風防をマイクロホンアレイに取り付けた状態で地表材質推定が可能であるかを判断するため、風防の性能を検証する実験を行うこととした。具体的には、外来ノイズをどの程度低減できるのか、目的音の抽出が可能であるかどうかの 2 点に焦点を当てた。詳細は 5.2 節にて説明する。

## 5 評価実験

### 5.1 地表材質推定

#### 5.1.1 利用するデータ

前述の通り、収録データには不具合のある部分が多いため、学習と推定に適さない区間を除去する。具体的には、以下の通りである。

- 材質を土砂、アスファルト、浅い水の 3 種類に限定
- 高度は 5m のデータのみを用いる
- 上記のうち、音割れが多い区間を除去（主に、浅い水の冒頭と終盤を除去し、8 分中 6 分 17 秒を抽出）
- 16 チャンネルのうち、データ値 0 区間が存在するデータを除去（その結果第 1, 4, 9, 10, 12, 13, 14, 15 チャンネルを選択）

なお、以降、浅い水を単に「水」と呼称する。



(a) 風防の各パーツ



(b) アレイに取り付けた状態

図 6: 提案したマイクロホンアレイ用風防

#### 5.1.2 実験条件

深層学習による分類を行うにあたり、以前の研究 [11] で利用された、図 7 に描く 4 層の畳み込みニューラルネットワーク (CNN) と、ResNet18 [15] の 2 つのネットワークを分類器として採用した。ネットワークの入力データを作成するために、各録音データを学習用、検証用、テスト用の 3 つの区間に分割した。全てのセクションに対し、表 2 に示す条件で短時間フーリエ変換 (STFT) を行い、スペクトログラムに変換した。直流成分を取り除いた後、データ長 256 フレーム、シフト長 128 フレームでスペクトログラムを切り出し、切り出した各スペクトログラムの実部と虚部を入力として利用した。

本実験では、8 チャンネルを使用した結果、学習用には  $16 \times 256 \times 256$  のサイズの特徴量が 1,094 個、検証用とテスト用にはそれぞれ 106 個得られた。学習データの教師ラベルは、地表材質に対応する 3 次元の one hot ベクトルを用いた。

地表材質ごとの各セクションの長さを表 3、CNN と ResNet18 の学習条件を表 4 に示す。学習曲線をもとに、過学習が生じない範囲で最適な性能のエポックを選択し、以下の式で示す正答率 (Accuracy) を用いて評価を行った。

$$\text{正答率 (\%)} = \frac{\text{推定ラベルと正解ラベルが一致したデータ数}}{\text{入力データ数}} \times 100$$

また、混同行列を計算し、材質ごとの傾向を分析した。

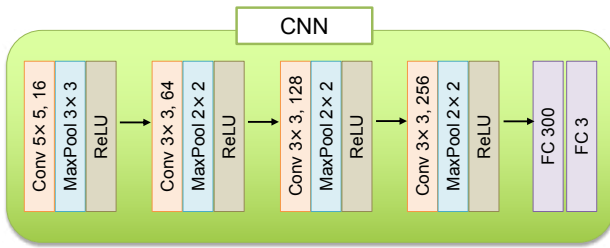


図 7: 本実験で用いる CNN アーキテクチャ

表 2: STFT 条件

STFT フレーム長	512
STFT シフト長	128
STFT 窓関数	ハン窓

### 5.1.3 実験結果

表 5 に各データに対する正答率を示す。テストデータの正答率は、CNN では、24 エポック終了時のモデルを使用し 70.8% であった。また、ResNet18 では、38 エポック終了時のモデルを使用し、82.1% となった。混同行列のヒートマップを図 8 に示す。

### 5.1.4 考察

CNN のテストデータに対する正答率は 70% と高い値を示しているが、検証データに対する正答率は、38% にとどまっている。本稿で扱ったタスクは 3 クラス識別であり、正答率のチャンスレートは、33% であることを考慮すると、この結果を額面通り捉えることはできず、原因の分析が必要である。学習過程で何らかの不具合が起きていた、データに偏りがあったといった理由が考えられるので、データをシャッフルする交差検定を行って原因の特定を行う予定である。一方、ResNet18 はいずれのデータに対しても 80% を超える正答率が得られており、地表材質推定の学習が適切に行えていると考えられる。CNN よりも ResNet18 の学習がよい理由は、CNN に比べ層数が多いため、これらの層で、反射音以外の雑音を抑圧する学習が適切に行われているからであろうと推察している。

混同行列のヒートマップを確認すると、いずれのネットワークもアスファルトと土砂に関しては再現率が高い反面、水に対する再現度が低い。特に CNN では、水と推定できたデータが 1 つもなく、アスファルトとの区別ができていないことがわかる。この原因として、ドローンのプロペラから生じる風は動的に変化するため、水面も定常的な地表面でなくなり、反射音の特徴が不

表 3: 各セクションの時間長 (単位: 秒)

地表材質	学習用	検証用	テスト用
アスファルト	405	40	40
土砂	407	40	40
水	313	32	32

表 4: 学習条件

損失関数	クロスエントロピー
オプティマイザ	Adam
学習率	$10^{-4}$
エポック数	100

規則に変化することがあげられる。しかし、現時点では仮説の域を出ないので、ネットワークごとの性能差とあわせて、今後分析を行う予定である。

## 5.2 風防の性能評価

本稿で提案したマイクロホンアレイ用の風防について、強風環境下での性能を確かめるべく、以下に示す実験を行った。

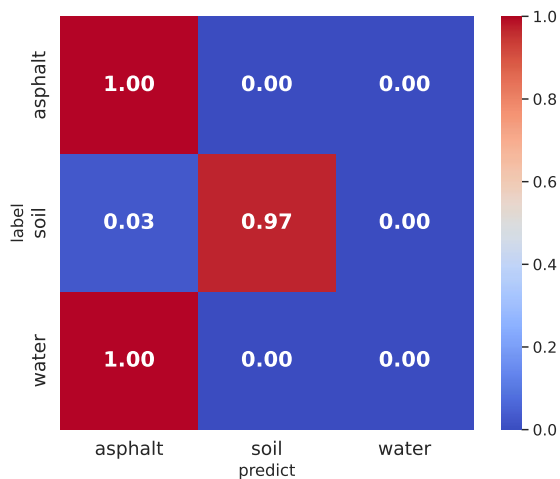
### 5.2.1 実験条件

実験を行うフィールド図を図 9 に示す。図中の緑色の線で示す、一周約 350m の周回路を走行する車内から、マイクロホンアレイを取り付けた治具と風速計を窓から押し出し、赤い丸で示す位置から発せられる目標音（汽笛音と音声）を収録した。このとき、治具が回転しないように制御しながら、マイクロホンアレイが車の窓から約 1m 離れた位置で保持した。

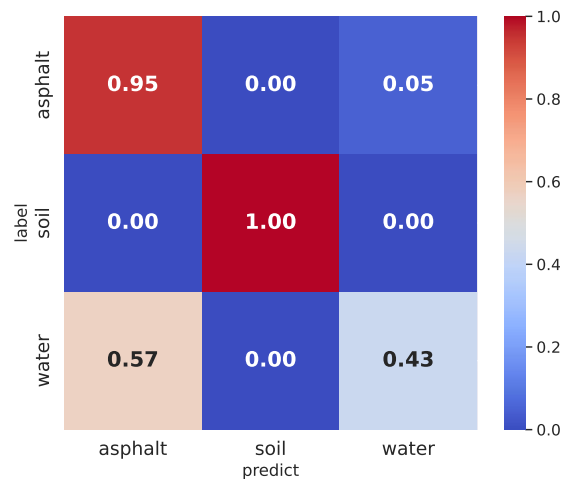
提案した風防の性能を検証すべく、以下の 4 つのマイクロホンアレイを収録に用いた。

1. 風防を何も取り付けないもの (bared)
2. 本稿で提案した風防を取り付けたもの (shield)
3. 風上側に位置するチャンネルに 10mm 厚のポリウレタンフォームを貼り付けたもの (PU)
4. 風上側に位置するチャンネルに 10mm 厚のポリエチレンフォームを貼り付けたもの (PE)

いずれの収録においても風速計の値が 12–20m/s となるよう周回を行った。周回路は私有地内に位置しており、各条件ごとに 3 周から 6 周走行した。評価指標には、それぞれの収録でのスペクトルと単位時間当たりのクリッピング発生率の 2 つを用いた。



(a) CNN の場合



(b) ResNet18 の場合

図 8: 混同行列のヒートマップ

表 5: 各データに対する正答率

ネットワーク	学習データ	検証データ	テストデータ
CNN	68.1%	36.8%	70.8%
ResNet18	100%	70.8%	82.1%

表 6: 1 秒間あたりの絶対値が  $th$  以上の観測点の個数

閾値	bared	shield (提案手法)	PU	PE
$th = 0.99$	16.59	0	0.358	0.682
$th = 0.9$	41.73	0	1.433	2.317

### 5.2.2 結果と考察

まず図 10 に各収録のスペクトルを示す。今回提案した風防はほぼすべての周波数で 10–20dB 程度の遮音効果があり、ほかのフォームと比較しても最も遮音効果が高いことが見て取れる。特に 3000Hz 以上の高周波領域では遮音性能が極めて高くなっており、今回のターゲット音が 100–3000Hz 程度であることと、風切り音には高周波成分が多く含まれていることを考慮すると、強風による音割れを防ぐのに今回提案した風防が有効であることが予想される。

次に、録音された観測値の絶対値が指定された閾値 ( $th$ ) 以上となる観測点の個数を表 6 に示す。ほかのフォーム素材では音割れを防ぎきれていないのに対し、提案した風防では音割れを示唆するデータ点が一つも存在しない。さらに、最大の観測値の絶対値も約 0.67 と、これ以上の強風にも耐えられる可能性を表している。したがって、先ほど示したスペクトルと合わせて、提案した風防が強風に極めて有効であることが示された。



図 9: 風防性能評価実験のフィールド図

## 6 おわりに

本稿では、ドローンノイズの反射音から頑健な地表推定手法を提示した。そのために、屋外環境下においてローターノイズと地表材質 4 種類（土砂、アスファルト、深さの異なる水 2 種類）に関するデータセットを作成した。このデータセットをもとに、深い水を除いた 3 種類を対象に、CNN および ResNet18 ベースの地表材質推定手法を構築した。また、生成したデータセッ

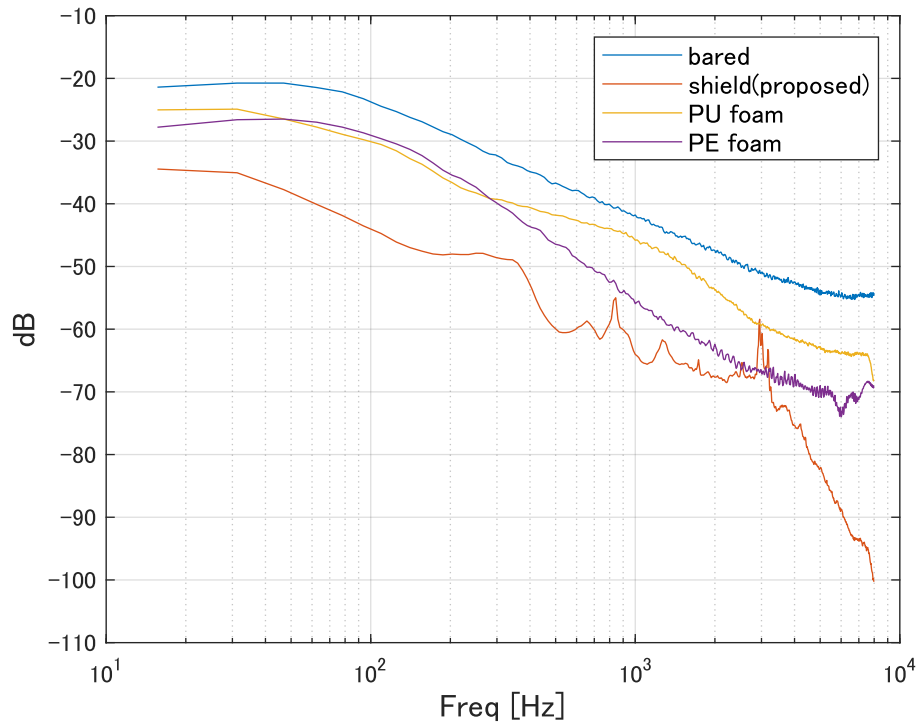


図 10: 各マイクロホンアレイの収録音スペクトル

トを用いて、構築した手法を評価した結果、ResNet18で正答率 70.8% とその有効性を示すことができた。さらに、外来ノイズへの対策としてマイクロホンアレイ用の風防を新しく提案し、強風による音割れを防ぐ効果があることを示すことができた。本データセットの拡張、今回提案した風防を用いてのデータ収録、ネットワークの違いに伴う識別容易性の分析、リアルタイム処理については、今後の課題である。

## 謝辞

本研究は JSPS 科研費 JP22F22769, JP22KF0141, および福島国際研究教育機構 (F-REI) の委託研究費 (JPFR23010102) の助成を受けた。

## 参考文献

[1] 国土交通省 砂防部：令和元年台風第 19 号に伴う土砂災害の概要 ver2.1, [https://www.mlit.go.jp/river/sabo//jirei/r1dosha/r1typhoon19\\_gaiyou191224r.pdf](https://www.mlit.go.jp/river/sabo//jirei/r1dosha/r1typhoon19_gaiyou191224r.pdf)

[2] 石川県：令和 6 年 (2024 年) 能登半島地震に関する情報 被害等の情報について (第 12 報), <https://www.pref.ishikawa.lg.jp/saigai/documents/202401041500higaihou.pdf>

[3] World Meteorological Organization: WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970–2019), <https://library.wmo.int/idurl/4/57564>

[4] 松井開, 長谷川克也, 大谷淳, 加藤洋一, 横澤真子: ドローン搭載カメラからの RGBD 画像に対するセグメンテーション等による土砂災害箇所と規模の推定法の検討, 研究報告コンピュータビジョンとイメージメディア (CVIM), No. 16, pp. 1–6 (2023)

[5] K. Okutani, T. Yoshida, K. Nakamura and K. Nakadai: “Outdoor auditory scene analysis using a moving microphone array embedded in a quadcopter”, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3288–3293, 2012

- [6] K. Hoshiba, K. Washizaki, M. Wakabayashi, T. Ishiki, M. Kumon, Y. Bando, D. Gabriel, K. Nakadai, and H. G. Okuno : “Design of UAV-Embedded Microphone Array System for Sound Source Localization in Outdoor Environments”, *Sensors*, vol. 17, no. 11, 2017
- [7] T. Yamada, K. Itoyama, K. Nishida and K. Nakadai : “Placement Planning for Sound Source Tracking in Active Drone Audition”, *Drones*, vol. 7, no. 7, 2023
- [8] M. Kumon, H. G. Okuno and S. Tajima : “Alternating Drive-and-Glide Flight Navigation of a Kiteplane for Sound Source Position Estimation”, *IEEE International Conference on Intelligent Robots and Systems (IROS)*, pp. 2114-2120, 2021
- [9] B. Yen, T. Yamada, K. Itoyama and K. Nakadai : “A Performance Assessment on Rotor Noise-Informed Active Multidrone Sound Source Tracking Methods”, *Drones*, vol. 8, no. 6, 2024
- [10] L. Wang, M. Clayton and A. G. Rossberg: “Drone audition for bioacoustic monitoring”, *Methods in Ecology and Evolution*, vol. 14, No. 12, pp. 3068-3082, 2023
- [11] 矢野翼, 糸山克寿, 西田健次, 中臺一博 : “ドローンのローターノイズによる地表材質推定手法の検討”, 第 24 回計測自動制御学会システムインテグレーション部門講演会 (SICE SI) , 2023
- [12] 江本久雄, 馬場那仰, 浅野寛元, 長瀬大和 : “AI手法による打音検査の浮き判定の検討”, *AI・データサイエンス論文集*, vol. 1, no. J1, pp. 514-521, 2020
- [13] 上半文昭 : “ドローンを用いたコンクリート橋部材の詳細検査手法”, *コンクリート工学*, vol. 57, no. 9, pp. 699-704, 2019
- [14] S. Dong, Y. Duan, X. Chen, F. You, X. Jiang, D. Wang, D. Hu, and P. Zhao : “Recent Advances in Preparation and Structure of Polyurethane Porous Materials for Sound Absorbing Application”, *Macromolecular Rapid Communications*, vol. 45, pp. 2400108, 2024
- [15] K. He, X. Zhang, S. Ren and J. Sun: “Deep Residual Learning for Image Recognition”, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016