アクティブ音響センシングによる ポイ捨てゴミの種別認識手法の提案

立花 巧樹[1]　松田 裕貴[1]　磯部 海斗[2]　真弓 大輝[1]　諏訪 博彦[1]　安本 慶一[1]　村尾 和哉[2]

概要：ゴミのポイ捨ては野生動物の殺傷に繋がるなど深刻な環境問題に発展している．ポイ捨てを防止する策を講じるために必要な，ポイ捨てゴミの実態やゴミ拾い活動の実績は情報として整理されていないのが実態である．したがって本研究では，ゴミの種別・位置情報を網羅的に収集するシステムの実現を目的としている．本研究の目的を達成するために，これまでに腕に装着しているスマートウォッチの音響センサを用いたポイ捨てゴミの認識手法や，ゴミ拾い用のトングに搭載した小型カメラを用いたポイ捨てゴミの認識手法を提案してきた．しかし，これらの研究では，ゴミの種別推定における範囲の制限，認識精度の精度不足，プライバシーの問題などが存在した．上記問題を解決するために本稿では，ゴミ拾い用トングにアクティブ音響センシングを組み合わせた，ゴミ種別認識システム Tongaraas を提案する．評価実験にて，缶・タバコ・紙類・プラスチック袋・ペットボトル・何も掴んでいない状態の 6 種類のデータセットを構築し，ゴミ種別認識モデルを構築した結果，F 値 0.622 が得られ，アクティブ音響センシングを用いたゴミ種別認識の実現可能性が示唆された．また，ゴミの形状や大きさによって，トングとゴミの接触の仕方や音響特性が大きく変化することが，認識性能に影響を与える可能性を明らかにした．

A Littering Garbage Recognition Method utilizing Active Acoustic Sensing

KOKI TACHIBANA[1]　YUKI MATSUDA[1]　KAITO ISOBE[2]　DAIKI MAYUMI[1]　HIROHIKO SUWA[1]
KEIICHI YASUMOTO[1]　KAZUYA MURAO[2]

1. はじめに

奈良公園で野生の鹿がポイ捨てされたゴミを食べて死に至る状況が問題となっている．この事例に留まらずポイ捨ては世界規模の問題であり，野生動物の健康状態に悪影響を与え，結果死に至らしめる要因の一つとなっている．この問題の解決は急務である．ポイ捨て問題を解決するためには，(1) 継続的なポイ捨てゴミの清掃，(2) ゴミ箱・ポイ捨て禁止を促す看板等の設置，(3) 行政の介入，という 3 つのアプローチがあると山根ら [1] は指摘している．しかしながら，これらを効果的に実施するために必要な，ポイ捨てゴミの実態やゴミ拾い活動の実績は情報として整理されていないのが実態である．結果として，ポイ捨てを未然に防ぐ施策は経験則に頼るものとなっており，またゴミ拾いの活動も地域やコミュニティで個別に実施されており連携がとれていない．こうした状況から，有料でポイ捨てゴミの分布を調査するサービスが存在し [2]，実際に複数自治体で運用されている．しかしながら，こういった調査は専用の調査人員が地域を巡回することを要し，時空間的な情報の網羅性や持続可能性に難がある．より簡便かつ負担の少ないポイ捨てゴミの情報収集技術が求められている．

本研究では，空間的・時間的に網羅性の高いゴミ種別・位置の情報収集を可能とするセンシング技術の実現を目指している．これまでの研究では，参加型センシングの枠組み [3] を日常的にゴミ拾いをしている人々に適用するアプローチを模索してきた．具体的には，腕に装着しているスマートウォッチの音響センサを用いたポイ捨てゴミの認識手法 [4] や，ゴミ拾い用のトングに搭載した小型カメラを用いたポイ捨てゴミの認識手法 [5] を提案している．しかしながら，前者の手法ではゴミを手で叩き音を発生させる

1 奈良先端科学技術大学院大学, Nara Institute of Science and Technology
2 立命館大学, Ritsumeikan University

必要があることや，叩いても音が発生しない対象物が認識できないという問題があり，後者の手法についてもゴミ拾いが実施されることの多い早朝といった周囲が暗い時間帯において画像認識精度が低下する傾向にあることがわかっている．

そこで，新たなポイ捨てゴミの認識を実現するアプローチとして「アクティブ音響センシング」に着目した．アクティブ音響センシングとは，物体表面にマイクとスピーカを取り付け，特定の音響信号をスピーカにより発し物体内部に伝搬させ，マイクで取得した応答信号を周波数解析をすることで物体の状態を推定する手法である．この手法をカトラリーに適用した事例では，接触した食材を認識することが可能であることが示されている [6]．本稿では，アクティブ音響センシングをゴミ拾い用トングに適用することで，ゴミ種別を認識するシステム Tongaraas を提案する．

評価実験では，缶・タバコ・紙類・プラスチック袋・ペットボトル・何も掴んでいない状態の 6 種類のデータセットを構築し，ゴミ種別認識モデルを構築した．その結果，F 値 0.622 での認識が可能であることが示され，アクティブ音響センシングを用いたゴミ種別認識の実現可能性が示唆された．しかしながら，ゴミの形状や大きさによって，トングとゴミの接触の仕方や音響特性が大きく変化することが，認識性能に影響を与える可能性を明らかとなった．加えて，異なる形状であったとしても似た素材で構成されるゴミ（プラスチック袋とペットボトル）の誤認識が発生することを確認した．

2. 関連研究

2.1 ポイ捨てゴミ識別の研究

ポイ捨てゴミを識別する研究は多く存在する．Huynh, Ahmad，Borge らの研究は，固定カメラで映像を収集し，収集した動画データを用いてゴミを機械学習で識別している [7], [8], [9]．しかし，固定カメラでは，カメラに映るゴミの種別・位置情報しか収集することができないため，ゴミの情報を広範囲で収集できないという問題が存在する．ゴミの情報を収集したい場所全体に固定カメラを配置することも考えられるが，設置台数などを考慮すると現実的でない．

井上ら [10] は，釜口水門から横川川合流点までの天竜川上流 11.5 km の河岸において，ポイ捨てゴミ手動で識別しながら調査し，漂着ゴミの分布および岸形状との関係を考察した．株式会社ピリカは，ゴミ拾いボランティアのための SNS アプリケーション，ピリカ [11] を開発している．ユーザは，ポイ捨てされたゴミをスマートフォンで撮影してアップロードすることができる．しかし，手動でポイ捨てゴミの種別・位置情報を記録する手法 [10] だとゴミの情報を入力するのに手間がかかり，効率的にゴミの種別・位置情報を収集できないという問題が存在する．同様にピリカ [11] を用いて，個々のポイ捨てゴミを写真でアップロードするのはユーザの手間がかかってしまうという問題が存在する．

株式会社ピリカは他にも，計測者がスマホで道に落ちているゴミを撮影すると，動画に写ったゴミの種類や数量，位置情報などを解析し，地図上にプロットするサービスを開発している [2]．Hong，Fulton，Kraft らは，ロボットのカメラ映像を用いて，機械学習でポイ捨てゴミを識別し，その地点の位置情報を収集している [12], [13], [14]．これらの研究は，ゴミの種別・位置情報を収集するだけでなく，実際にゴミを拾うことまで想定している．しかし，ゴミの情報を収集したい場所でロボットを運用するためのコストや，ゴミの情報を調査するための人件費が必要になるため，運用コストが高くなるという問題が存在する．

上記問題点を解決するため，我々はこれまでに，腕に装着しているスマートウォッチの音響センサを用いたポイ捨てゴミ種別認識システム ACOGARE [4] を提案してきた．しかしながら，ACOGARE ではゴミを手で叩き音を発生させる必要があることや叩いても音が発生しない対象物が認識できないという問題がある．例えば，タバコを叩いても音が発生しないことは既存の知見から把握している [15]．他にも，ゴミ拾い用のトングに搭載した小型カメラを用いたポイ捨てゴミ種別認識システム Tongar [5] を提案してきた．これにより，タバコなどのゴミも認識可能になったが，この Tongar についてもゴミ拾いが実施されることの多い早朝といった周囲が暗い時間帯において画像認識精度が低下する傾向にあることがわかっている．また，スマートウォッチ内蔵マイクから収集する音響データやトング搭載カメラによって収集する画像データから，プライバシーの問題が発生することが懸念される．

2.2 アクティブ音響センシングによる物体識別の研究

物体表面にマイクとスピーカを配し，特定の音響信号を物体内部に伝搬させ，その応答の周波数解析をすることで状態を推定する手法であるアクティブ音響センシングを用いる物体識別の研究は多く存在する．

大野らはこのアクティブ音響センシングを用いたタッチ認識技術を提案している [16], [17]．これは，既存物体にコンタクトマイクとコンタクトスピーカを取り付け，物体の触り方の違いを取得することで把持状態を認識する手法である．河田らは，この把持状態の認識を，物体ではなくユーザにセンサを取り付けることにより行った [18]．また，Adiyan らは低周波の超音波で皮膚表面を共振させ，身体上の別地点で受け取る．これにより，身体上の接触位置やジェスチャを推測できる [19]．Kim らは，携帯電話に内蔵されているスピーカの音と振動モーターからの振動を利用して，加速度センサとマイクを用いて，応答を分析することで携帯電話の位置を把握している [20]．また，岩瀬ら

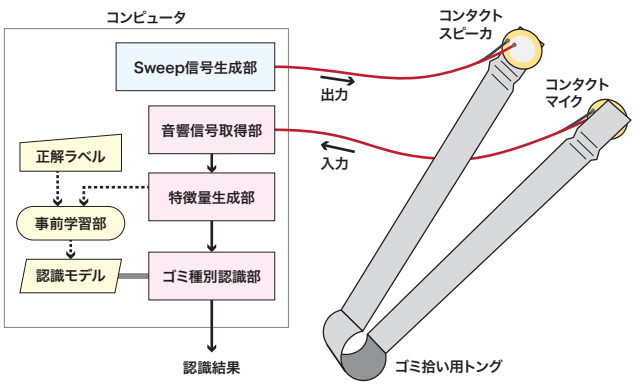

図 1　Tongaraas のシステム概要

は，日常生活を支援するために，実空間内にある物体情報を取得するための物体情報識手法を提案している [21]．また，西井らは，アクティブ音響センシングが可能なフォーク型・スプーン型デバイスを開発し，接触した食材を認識する手法を提案している [6]．

このように，アクティブ音響センシングを用いることで接触した物体を認識することが可能であることが示されており，その技術は様々な用途に活用されている．

2.3　本研究の立ち位置

本研究では，これまでの研究（ACOGARE [4], Tongar [5]）で提案してきた，ゴミを拾うだけでゴミの種別・位置情報を収集可能な参加型センシングシステムをベースとしつつ，アクティブ音響センシングをゴミ認識の新たなアプローチとして取り入れることで，認識可能な対象範囲の拡大や環境依存性の低減を目指す．

3.　提案手法

本章では，ゴミ拾い用トングにアクティブ音響センシングを取り入れた，新たなゴミ種別認識システム Tongaraas（IoT Tongs for Littering Garbage Recognition with Active Acoustic Sensing）を提案する．

3.1　概要

ゴミ拾い用トングを用いたゴミ拾いでは，収拾時に認識の対象となるゴミとトングが接触することや，認識対象物が異なる音響特性を持つ素材（金属，紙，プラスチックなど）で構成されている．例えば，缶やペットボトルは円柱状になっているのに対し，プラスチック袋は袋状になっている．こうした形状の違いが周波数特性に現れると考える．また，ゴミの種別ごとの材質も同様に異なる．缶は主にアルミやスチールで製造されており，ペットボトルはポリエチレンテレフタレートで製造されている．この材質の違いが同様に音の周波数特性に表れると考えている．以上から，アクティブ音響センシングは，ゴミの種別を推定するタスクに適応できると考えた．

本システムでは，タバコなどの ACOGARE で分類できなかったゴミの種別を分類可能になること，また，周囲が暗い時間帯において画像認識の精度が低下する Tongar の問題を低減することが可能になると考えられる．また，超音波領域の音響信号を用いたセンシングを用いるため，可聴域の音データを用いる ACOGARE や映像データを用いる Tongar とは異なりプライバシーの侵害が起きづらいという利点がある．これにより，提案手法でゴミの種別をセンシングすることができれば，認識精度の向上やゴミの種別の範囲が拡張され，プライバシーの保護が期待される．

3.2　システム構成とゴミ種別認識の手順

Tongaraas のシステム概要を**図 1** に示す．システムは，音響信号を発するコンタクトスピーカと音響信号を収集するコンタクトマイクを装着したゴミ拾い用トング，Sweep 信号生成部，音響信号取得部，特徴量生成部，ゴミ種別認識部からなるコンピュータにより構成される．

ゴミ種別認識の手順は次のとおりである．

(1) 周波数が時間とともに変化する Sweep 信号をコンピュータの Sweep 信号生成部で生成，コンタクトスピーカから発する．
(2) トングを伝播し伝わった音響信号をコンタクトマイクによって収集する（この際，接触しているゴミの種類や形状によって音響信号に変化が生じる）．
(3) 収集した音響データから Sweep 信号の周波数帯域を取り出し，特徴量抽出を行う．
(4) 事前に学習したゴミ種別認識モデルを用いて，トングに接触している未知のゴミの種別認識を行う．

次節以降では，Sweep 信号生成部・音響信号収集部，および特徴量抽出部・ゴミ種別認識部について詳述する．

3.3　Sweep 信号生成部・音響信号収集部

Sweep 信号生成部では，周波数が 20kHz から 40kHz まで直線的に変化する Sweep 信号（Chirp 信号とも呼ばれる）を生成する．この周波数帯域は，先行研究 [16] の論文を参考に決定した．Sweep 信号を繰り返し再生する際には，繰り返しの切れ目（40kHz まで変化したのち，20kHz に戻る瞬間）に周波数が大きく変化することで発生するインパルスノイズによって，認識に影響が生じることが考えられる．そこで本システムでは，**図 2** に示すように，次の 200ms で周波数を増加（20→40kHz）させた後，その次の 200ms で周波数を減少（40→20kHz）させる．合わせて 400ms 周期の Sweep 信号を生成することでこの問題を解消する．なお，音響信号の収集にあたっては，トングを介して伝わったアナログ Sweep 信号をデジタル信号に変換するため，サンプリング周波数は 96kHz とする．

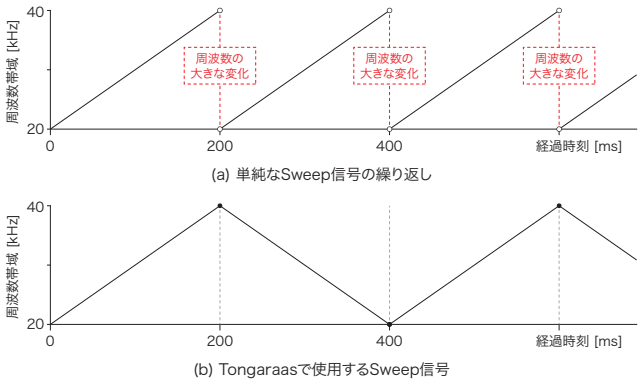

図 2　Sweep 信号の生成方法

表 1　特徴量一覧

特徴量	次元数
MFCC	104
Chroma	12
Melspectrogram	128
Spectral contrast	7
合計	251

3.4　特徴量抽出部・ゴミ種別認識部

事前処理として，Sweep 信号とは異なる周波数帯域の音響信号の影響を排除するため，収集した音響データからバンドパスフィルタを用いて 20kHz から 40kHz の周波数帯を抽出する．

次に，音響データから特徴量を抽出する．抽出する特徴量の一覧は表 1 に示すとおりであり，その次元数は 251 である．本システムでは，各特徴量を Python の librosa ライブラリ[*1]を用いて取得する．これらの特徴は，音響分析において一般的に利用されている特徴量であり，MFCC は，音データに対して FFT を行い，メルフィルタバンクを通した出力を逆離散コサイン変換した特徴量，Chroma は，オクターブ違いの同じ音階の成分を全て重ね合わせて 1 オクターブ内の半音階の 12 音の成分に縮約した特徴量，Mel spectrogram は，FFT 変換を行った後にスペクトログラムを作成し，周波数がメルスケールに変換した特徴量，Spectral contrast は，FFT を行い，オクターブフィルタバンクを通した後にピークを検知して抽出し，最後にカルーネン・レーベ変換した特徴量である [22].

最後に，事前に取得した特徴量とゴミの正解ラベルからなる学習データを用いてゴミ種別認識モデル（本稿での評価では LightGBM を使用）を構築しておき，未知のゴミを認識する．学習データには認識したいゴミの特徴量に加えて，トングがゴミに触れていないときのサンプルを含めている．これによりトングでゴミを把持しているときはゴミの種別が認識され，それ以外のときには触れていないと判定される．

*1　https://github.com/librosa/librosa

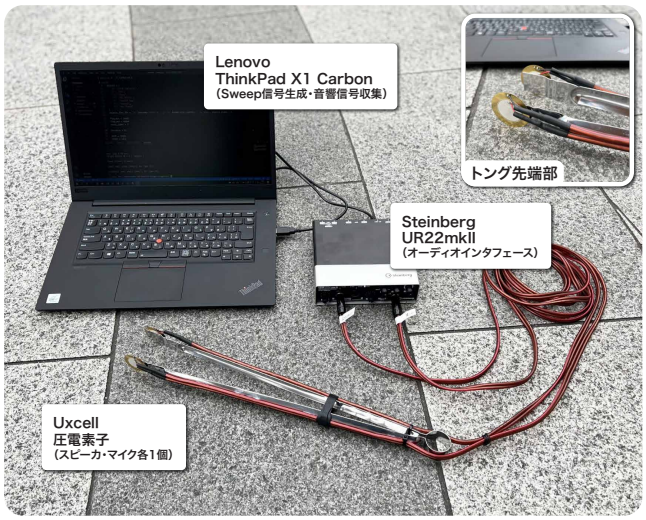

図 3　Tongaraas のデータ収集実験セットアップ

表 2　分類ラベルの一覧

ラベル名	ゴミの種類
Can	缶（アルミ缶）
Tobacco	タバコ
Paper	紙類
Plastic_bag	プラスチック袋
Pet_bottle	ペットボトル
None	何も掴んでいない

4.　評価実験

提案手法の有効性を検証するため，種類ごとに複数個のゴミを準備しデータ収集・ゴミ種別認識モデル構築を行い評価する．

4.1　システムセットアップ

評価実験に用いるデータ収集システムのセットアップを**図 3** に示す．Sweep 信号生成，音響信号収集を行うコンピュータには Lenovo ThinkPad X1 Extreme Gen3（CPU：Intel Core i7-10750H，RAM：64GB，OS：Windows 10）を使用し，音の入出力は USB 接続のオーディオインタフェース Steinberg UR22mk を介して行う．使用したコンタクトスピーカ・マイクとしては，uxcell 社製圧電素子（直径 35mm，厚さ 0.3mm）を使用し，図 3 右上に示すようにトングの先端部に取り付けた．

4.2　認識対象

本評価においては，ポイ捨てゴミの代表的なものとして，缶・タバコ・紙類・プラスチック袋・ペットボトルの 5 種類，加えて，ゴミを把持していない状態の計 6 種類を対象とした（**表 2**）．各ゴミ種別につき 10 個のゴミを用いた．各ゴミの外観は**図 4** に示すとおりである．なお，同じ種類のゴミであっても素材や形状が大きく異なるものが含

図 4　評価実験に用いたゴミの写真（5 種・各 10 個）

表 3　実験結果

	Leave-one-garbage-out（LOGO）交差検証の試行回ごとの結果										平均
	1 回目	2 回目	3 回目	4 回目	5 回目	6 回目	7 回目	8 回目	9 回目	10 回目	
Precision	0.313	0.754	0.861	0.823	0.685	0.404	0.751	0.686	0.589	0.745	0.661
Recall	0.364	0.711	0.808	0.794	0.647	0.408	0.808	0.672	0.569	0.722	0.651
F 値	0.333	0.665	0.746	0.774	0.663	0.404	0.761	0.663	0.540	0.675	0.622

まれている．例えば，缶については容量（185mℓ，300mℓ，355mℓ など）や，形状（プルタブ式・キャップ式，凹み度合い）が大きく異なっている．

4.3　データ収集・評価手順

4.1 節にて述べたシステムを用いて，4.2 節のゴミのデータ収集を行う．手に持ったトングで対象のゴミを把持した状態で，Sweep 信号を発するとともに音響信号を録音する．各ゴミにつき 62 秒分の音響データを収集した．次に，開始終了時の音響ノイズを除外するため，開始終了の 1 秒を削除した 60 秒分の音響データについて，1 秒ごとに分割し 60 サンプルの音響ファイルを生成する．これにより，合計 3600 サンプル（6 種類 ×10 個 ×60 サンプル）の音響データを生成する．

ゴミ種別認識モデルの構築にあたっては，音響データからゴミの種別を推定するタスクにおいて，他の機械学習の手法に比べ，分類精度が高くなることがわかっているモデル [4] として，LightGBM を用いた．また，ハイパーパラメータのチューニングは Hyperopt を用いた．これにより，251 次元の特徴量（表 1）を入力とし，6 種類のゴミ種別（表 2）を出力とするゴミ種別認識モデルを構築する．

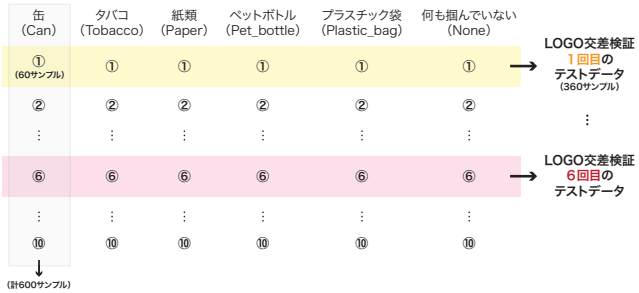

図 5　LOGO 交差検証の学習データ・テストデータの分割方法（各番号は図 4 の番号と対応）

評価には，Leave-one-garbage-out（LOGO）交差検証を用いる．LOGO 交差検証における学習データ，テストデータの分割方法を**図 5** に示す．具体的には，4.3 節で述べた音響データ 3600 サンプルをゴミ種別ごとにひとつずつ抜き去ることで，学習データ 3240 個（6 種類 ×9 個 ×60 サンプル），テストデータ 360 個（6 種類 ×1 個 ×60 サンプル）に分けて交差検証を行う．つまり，学習データには同一の個体のデータを一切含まない設定で学習を行う（今回のデータセットにおいては，10 回の試行を行う）．

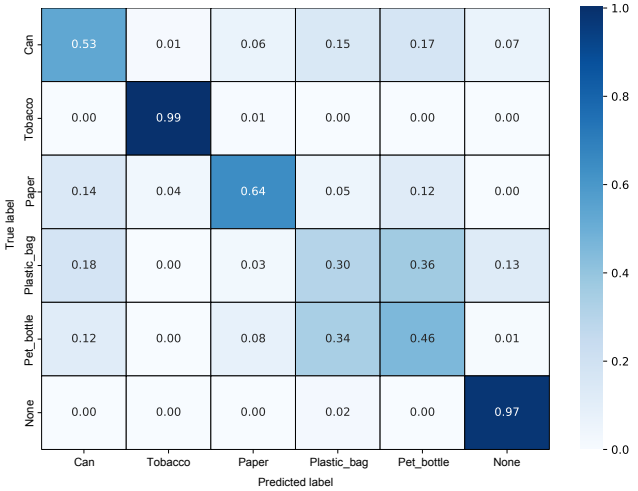

図 6　LOGO 交差検証の混同行列

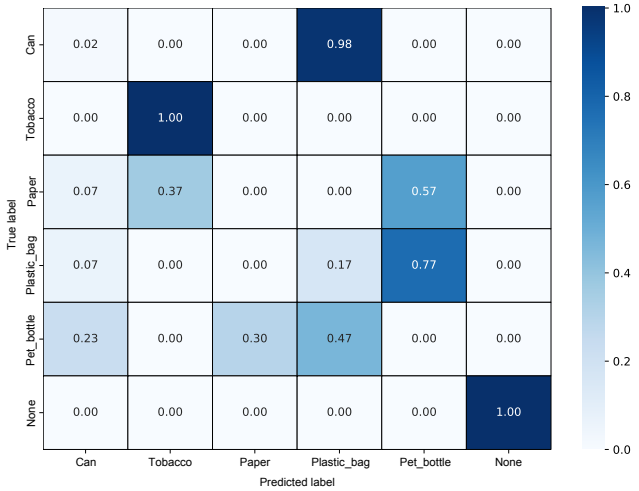

図 7　LOGO 交差検証 1 回目の試行における混同行列

5.　結果・考察

評価結果を**表 3** および**図 6** に示す．10 分割交差検証の結果，Precision の平均は 0.661，Recall の平均は 0.651，F 値は 0.622 となった．表 3 の LOGO 交差検証の試行回ごとの結果に注目すると，試行回によって認識性能が大きくばらつく結果となった．以降では，特徴的な結果を示していた LOGO 交差検証 1 回目および 7 回目の試行における混同行列（**図 7**，**図 8**）を用いて考察する．

LOGO 交差検証 1 回目（図 7）においては，Tobacco と None の精度は 100%であるが，それ以外のゴミの種別は極めて低くなっていることがわかる．特に Can は，ほぼ全て Plastic_bag と誤認識する結果となっている．これは，Can の凹み度合いによってトング先端の面に対象物が上手く接触していないことが原因の一つであると考えられる．また，Paper は Tobacco や Pet_bottle と誤認識している．これは，Paper の中で紙袋のサンプルが当該サンプルしかなかったため，未知の物体として取り扱われてしまったものと考えられる．対して，LOGO 交差検証 7 回目（図 7）では，Pet_bottle 以外のゴミにおいては全て 90%以上の精度を算出している．このことから学習データに類似の素材・大きさ・形状の物体が含まれている場合には正確な推定が可能である可能性が示唆されている．

加えて，Plastic_bag と Pet_bottle については，全体を通して互いに誤認識する傾向にある結果となった．これは，Pet_bottle と Plastic_bag の素材はどちらもプラスチックのため，誤認識が生じている可能性がある．今後は，Pet_bottle と Plastic_bag でも誤認識が生じないようなデータ収集方法やモデル構築方法について検討する必要がある．なお，今回の評価実験においては，1 つの個体については各 1 回のみデータ収集しているため，ゴミの把持方法のバリエーションが十分でなかった可能性がある．今後は，1 つのゴミに対し把持方法を変えてデータ収集するなど，汎化性能を高められるようなデータ収集を行う予定である．

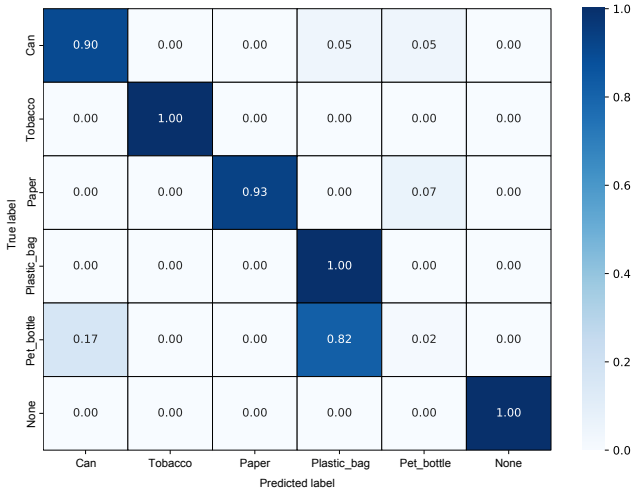

図 8　LOGO 交差検証 7 回目の試行における混同行列

6.　おわりに

本稿では，ゴミの種別・位置情報を網羅的に収集するシステムの実現を目的として，ゴミ拾い用トングにアクティブ音響センシングを組み合わせた，ゴミ種別認識システム Tongaraas を提案した．評価実験にて，6 種類のデータセットを用意し，ゴミ種別認識モデルを構築した結果，F 値 0.622 が得られ，アクティブ音響センシングを用いたゴミ種別認識の実現可能性が示唆された．しかしながら，ゴミの形状や大きさによって，トングとゴミの接触の仕方や音響特性が大きく変化してしまうことから，認識性能に影響が出ることや，形状が異なるものの似た素材で構成されるゴミの誤認識が確認された．今後は，異なる把持方法である場合や，素材が同じでゴミの種別が異なる場合でも，誤認識が生じないようなデータ収集方法やモデル構築方法について検討する必要がある．また，1 つのゴミに対し把持方法を変えてデータ収集するなど，汎化性能を高められるようなデータ収集を行う予定である．

謝辞

本研究の一部は，JST さきがけ JPMJPR2039，JPMJPR1937 の助成を受けて行われたものです.

参考文献

[1] 山根大路, 松村真宏. 見えざる人の存在を想起させる仕掛けによるポイ捨て抑止実験.

[2] 株式会社ピリカ. ポイ捨てごみ調査サービス「タカノメ」—スマホで簡単、ポイ捨て分布調査. https://research.pirika.org/. Accessed: 2020-07-27.

[3] J. A Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B Srivastava. Participatory sensing. *Center for Embedded Network Sensing*, 2006.

[4] 立花巧樹, 中村優吾, 松田裕貴, 諏訪博彦, 安本慶一. スマートウォッチの音響センサを用いたポイ捨てごみの種別認識手法の提案と評価. マルチメディア, 分散, 協調とモバイル（DICOMO 2021）シンポジウム論文集, pp. 325–332, 2021.

[5] 立花巧樹, 中岡黎, 宮地篤士, 冨田周作, 松田裕貴, 中村優吾, 諏訪博彦. センサ装着型トングを用いたポイ捨てごみの種別認識手法の提案と評価. 第 28 回社会情報システム学シンポジウム（ISS28）, pp. 1–6, 2022.

[6] 西井遥菜, 双見京介, 村尾和哉. アクティブ音響センシングを用いた食材認識カトラリーの 設計と評価. 一般社団法人情報処理学会 インタラクション 2021 論文集, pp. 164–167, 2021.

[7] M. Huynh, Phu-Thinh Pham-Hoai, Anh-Kiet Tran, and Thanh-Dat Nguyen. Automated waste sorting using convolutional neural network. *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, pp. 102–107, 2020.

[8] Kashif Ahmad, Khalil Khan, and Ala Al-Fuqaha. Intelligent fusion of deep features for improved waste classification. *IEEE Access*, Vol. 8, pp. 96495–96504, 2020.

[9] João Borges de Sousa, Ana Rebelo, and Jaime S. Cardoso. Automation of waste sorting with deep learning. *2019 XV Workshop de Visão Computacional (WVC)*, pp. 43–48, 2019.

[10] Yoshiki Inoue and Hideshige Toda. Drifted litters of lake suwa and upper reaches of the tenryu river. *ENVIRONMENTAL SCIENCE*, Vol. 16, No. 3, pp. 167–178, 2003.

[11] Pirika, Inc. Pirika — anti-litter social media. https://en.sns.pirika.org/. Accessed: 2021-07-09.

[12] Jungseok Hong, Michael Fulton, and Junaed Sattar. Trashcan: A semantically-segmented dataset towards visual detection of marine debris. *ArXiv*, Vol. abs/2007.08097, , 2020.

[13] Michael Fulton, Jungseok Hong, Md. Jahidul Islam, and Junaed Sattar. Robotic detection of marine litter using deep visual detection models. *2019 International Conference on Robotics and Automation (ICRA)*, pp. 5752–5758, 2019.

[14] Marek Kraft, Mateusz Piechocki, Bartosz Ptak, and Krzysztof Walas. Autonomous, onboard vision-based trash and litter detection in low altitude aerial images collected by an unmanned aerial vehicle. *Remote. Sens.*, Vol. 13, p. 965, 2021.

[15] 立花巧樹, 中村優吾, 松田裕貴, 諏訪博彦, 安本慶一. スマートウォッチを用いたポイ捨てごみの種別・位置認識システムの提案. 2020 年度 情報処理学会関西支部 支部大会 講演論文集, 第 2020 巻, 2020.

[16] 大野誠, 志築文太郎, 田中二郎. アクティブ音響センシングを用いた把持状態認識. 一般社団法人情報処理学会 インタラクション 2013 論文集, pp. 55–63, 2013.

[17] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. Touch & activate: Adding interactivity to existing objects using active acoustic sensing. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, p. 31–40, New York, NY, USA, 2013. Association for Computing Machinery.

[18] 河田大史, 寺田努, 塚本昌彦. 振動のアクティブセンシングによるウェアラブル把持物体認識システム. 情報処理学会マルチメディア，分散，協調とモバイルシンポジウム (DICOMO2015) 論文集, 第 2015 巻, pp. 400–406, 2015.

[19] Adiyan Mujibiya, Xiang Cao, Desney S. Tan, Dan Morris, Shwetak N. Patel, and Jun Rekimoto. The sound of touch: On-body touch and gesture sensing based on transdermal ultrasound propagation. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, p. 189–198, New York, NY, USA, 2013. Association for Computing Machinery.

[20] Kai Kunze and Paul Lukowicz. Symbolic object localization through active sampling of acceleration and sound signatures. Vol. 4717, pp. 163–180, 09 2007.

[21] 岩瀬大輝, 伊藤雄一, 秦秀彦, 山下真由, 尾上孝雄. アクティブ音響センシングによる日常物体識別と位置推定. 一般社団法人情報処理学会 インタラクション 2018 論文集, pp. 62–71, 2018.

[22] Dan-Ning Jiang, L. Lu, H. Zhang, J. Tao, and Lianhong Cai. Music type classification by spectral contrast feature. *Proceedings. IEEE International Conference on Multimedia and Expo*, Vol. 1, pp. 113–116 vol.1, 2002.