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

立花 巧樹[1] 中村 優吾[2] 松田 裕貴[1,3] 諏訪 博彦[1,4] 安本 慶一[1]

概要：ごみのポイ捨ては社会問題に発展している．ポイ捨てを防止する方法として，ごみ箱やポイ捨て禁止の旨が書かれた看板をポイ捨てされる場所に設置して，未然にポイ捨てを防止する都市デザイン的アプローチが考えられる．しかし，ごみ箱やポイ捨て禁止の旨が書かれた看板を，ポイ捨てが起こりうる場所全てに配置することは，現実的ではない．したがって，ポイ捨てが多くされる場所に対して効率的に配置することが求められる．そのためには，ごみの種別・位置を網羅的に事前に把握することが必要である．しかし，既存手法では，本問題を解決することは困難である．本研究では，ユーザが装着しているスマートウォッチから得られるセンサデータのみを用いて，ごみの種別および位置を推定するシステムを提案してきた [1]．本システムは，ユーザがごみを叩きながらごみ拾いをするのみで，ごみの種別・位置を記録することが可能である．本稿では，スマートウォッチの音響センサから得られるデータを用いてモデルの構築を行い，評価実験を行った．実験結果より，LightGBM を用いた際におけるモデルの精度が最も高くなった．

Evaluation of Classification and Location Recognition System for Litter by Using Smartwatch

KOKI TACHIBANA[1] YUGO NAKAMURA[2] YUKI MATSUDA[1,3] SUWA HIROHIKO[1,4]
KEIICHI YASUMOTO[1]

1. はじめに

ごみのポイ捨ては社会問題に発展している．Maria ら [2] によると，世界では，年間 6 兆本ものタバコが消費されているが，そのうち 4.5 兆本はポイ捨てされている．また，世界経済フォーラム [3] によると，2015 年にタバコを含む 910 万トンのプラスチックごみが流出したという調査結果を発表した．これにより，海洋汚染や海洋生物の殺傷などの問題が生じている．したがって，ごみのポイ捨ては早急に解決すべき問題である．ポイ捨てを防止する方法として，都市デザイン的アプローチが考えられる．例えば，ポイ捨てされる場所に対して，ごみ箱やポイ捨て禁止の旨が書かれた看板を設置することで，ポイ捨てを防止することができると考えられる．ただし，ごみ箱やポイ捨て禁止の旨が書かれた看板を，ポイ捨てが起こりうる場所全てに配置することは，現実的ではない．したがって，ポイ捨てが多くされる場所に対して効率的に配置することが求められる．例えば，タバコのごみが多く落ちている場所を事前に把握しておけば，その場所に灰皿を設置することで，未然にポイ捨てを防ぐことができる．そのためには，ごみの種別・位置を網羅的に把握することが必要である．

しかし，ごみの種別・位置を網羅的に把握するのは，手動でごみを記録する手法 [4] やビデオカメラなどでごみを自動で記録する手法 [5] では困難であるという問題が存在する．ごみ拾いをしている人が一つ一つのごみを撮影し，ピリカ [6] にアップロードすることは手間がかかってしまう．よって我々は，ごみを拾う人が，手間をかけずにごみを拾うだけで，ごみの種別・位置を網羅的に記録できるようにすることを研究の目的とする．

我々は，研究の目的を達成するアプローチとして，プロギングしている人に注目する．プロギングとは，スウェーデン発の新しいフィットネスであり，ジョギングしながら

1 奈良先端科学技術大学院大学
2 九州大学
3 JST さきがけ，JST PRESTO
4 理化学研究所 革新知能統合研究センター，RIKEN Center for Advanced Intelligence Project (AIP)

© 2021 Information Processing Society of Japan

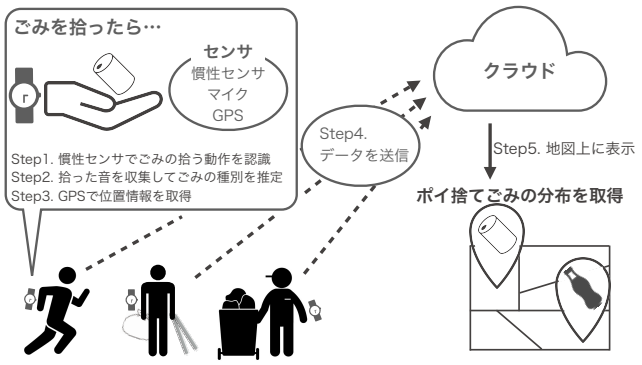

図 1　提案システムの概念図

ごみ拾いをすることである．我々は，プロギングをしている人が運動管理のために装着しているスマートウォッチを用いて，普段通りごみを拾うだけで個々のごみの種別・位置を記録することができれば，ユーザに手間をかけることなく網羅的にごみのデータ収集をすることができると考える．

本研究では，ユーザが装着しているスマートウォッチから得られるセンサデータのみを用いて，ごみの種別および位置を推定するシステムを提案してきた [1]．提案システムの概念を**図 1** に示す．本システムは，ユーザがごみを叩きながらごみ拾いをするのみで，ごみの種別・位置を記録することが可能である．具体的には，デバイスに内蔵されている IMU を用いて加速度データを収集することで，ユーザがごみを拾う動作を認識する（Step1）．つぎに，ユーザはごみを拾う際，ごみを叩く動作を行う．システムは，その際に生じる音をマイクで収集し，データを解析することでごみの種別を推定する（Step2）．最後に，GPS から得られる位置情報を取得することで（Step3），ごみの種別・位置を記録することができると考える．そして，ごみの種別と位置情報をクラウドに送信し（Step4），位置情報をもとに地図上にごみの種別をプロットすることで（Step5），ごみの分布を把握することが可能になると考える．本システムは，プロギングしている人を想定ユーザとしているが，ボランティアでごみ拾いをしている人やごみ収集業者がシステムを利用することも可能である．本稿では，スマートウォッチの音響センサから得られる音データを用いてモデルの構築を行い，様々な条件におけるモデルの精度を評価する．

2.　関連研究

2.1 節ではごみの種別・位置を手動で記録する事例，2.2 節ではごみの種別・位置を自動で記録する事例について述べる．

2.1　ごみの種別・位置を手動で記録する事例

ごみの種別・位置を手動で記録する事例として [4], [7], [8] がある．井上ら [4] は，釜口水門から横川川合流点までの天竜川上流 11.5km の河岸において，漂着ごみの量，種類を手動で記録しながら調査し，漂着ごみの分布および岸形状との関係を考察した．高橋ら [7] は，計測者が手動でタバコの位置を地図上にプロットし，タバコの種別・位置の記録を手動で行っている．その結果，タバコの多い場所に対して，灰皿を設置してポイ捨て削減を狙うなどのタバコのポイ捨ての防止対策に関する具体的方策を提案している．早瀬ら [8] は，計測者が大学敷地内に散乱しているごみの種別・位置を手動で記録している．また，ごみを捨てられた時間における滞留・通過人員も定点カメラで撮影することにより計測している．

しかし，これらの手法では，ごみ拾いをしている人が，一つ一つのごみに対し手動で記録するのは手間がかかるという問題が存在する．したがって，ごみの種別・位置を自動で検出することが必要である．

2.2　ごみの種別・位置を自動で記録する事例

ごみの種別・位置を自動で記録する事例として [5], [6], [9], [10] がある．タカノメ [9] は，スマホでごみを撮影することで，ごみの種別・位置を自動で把握できるシステムである．計測者がスマホで道にあるごみを撮影する．システムは，動画に写ったごみの種類や数量，位置などを解析し，地図上にプロットする．植田ら [5] は，ビデオカメラによるポイ捨て位置やポイ捨てしている人の特徴を記録している．ビデオカメラで撮影したデータ対して，タカノメの技術 [9] を用いれば，ごみの種別・位置を自動で記録することが可能になる．ピリカ [6] は，世界 100 ヶ国以上で利用されているごみ拾いボランティア SNS アプリケーションである．アプリケーションユーザは，拾ったごみを撮影し，位置情報と共にアップロードする．そして，アップロードされたごみの種別をアプリ管理者がタカノメ [9] の解析技術を利用することで，ごみの種別・位置を網羅的に把握することができる．Gong ら [10] は，スマホでモノを叩いて，その音から生活空間に存在するモノの種別を自動で推定する技術，Knocker を提案している．Knocker をごみの種別モデルに転移学習することによって，ユーザがスマホでごみを叩くだけで，システムがごみの種別・位置を自動で記録することが可能である．

しかし，文献 [5] の手法は，カメラが設置してある場所でしかごみを記録できないため，ごみを網羅的に記録することができないという問題が存在する．もし，多くの場所にカメラしている環境があったとしても，多くのコストがかかってしまう．上記の問題は，ごみを拾う人からデータを収集するアプローチ [6], [9], [10] であれば，モノの設置コストがかかることなく，ごみの種別・位置を網羅的に記録できるかもしれない．しかし，タカノメ [9] は，データ計測者がスマホで道路を撮影しながら歩くという動作が必

© 2021 Information Processing Society of Japan

要となってしまう．同様に，ピリカ [6] のアプリケーションユーザは，落ちているごみを一つ一つ撮影しなければならないため，手間が大きくなってしまう．Knocker[10] は，ピリカやタカノメ [6], [9] に比べてユーザの手間は少ないが，ごみをスマホで叩くのは現実的ではない．したがって我々は，ごみを拾う人が，手間をかけずにごみを拾うだけで，ごみの種別・位置を網羅的に記録できるようにすることを研究の目的とする．

本研究では，ユーザが装着しているスマートウォッチから得られるセンサデータのみを用いて，ごみの種別および位置を推定するシステムを提案してきた [1]．提案システムは，ごみを拾う人がごみを叩いた音をマイクで収集し，機械学習を用いてごみの種別を自動で記録する．ごみを叩くという動作は，ごみを撮影するという動作に比べ，手間が少ないと考えるため研究の目的を達成すると考える．本稿では，提案システム [1] のごみの種別を推定するモデルの構築手法を提案する．

3. モデル構築

3.1 モデル構築の手順

モデル構築の手順を**図 2** に示す．はじめに，ユーザがごみを叩く動作を行った際に，加速度からごみを叩いたことを機械学習で推定する（手順 1）．つぎに，ごみを叩く動作を行った時間の音を切り出し，前処理を行う（手順 2）．システムは，切り出した音から特徴量の抽出を行う（手順 3）．特徴量については，3.2 節で述べる．最後に抽出した特徴量を用いてモデル構築を行う（手順 4）．分類ラベルは，文献 [11] の散乱個数ランキングで上位を占めた，紙類・プラスチックに加え，一般的な飲料容器である缶・瓶・ペットボトルを加えて計 5 種類とした（**表 1**）．本実験では，手順２～４の順番でモデル構築を行った．また，タバコのような音がならないごみは，ユーザがスマートウォッチを 1 タップすることで，システムは，タバコだと認識することができるアプローチを検討している．また，システムが推定したごみの種別の確信度の低い場合，システムがユーザ自身にごみの種別を聞くことで，誤認識を防ぎ，ユーザの叩き方にモデルが適用していく設計にすることも検討している．

表 1 分類ラベルの一覧

教師ラベル	ごみの種類
Bottles	瓶
Can	缶
Paper	紙類
Plastic	プラスチック
Pet bottle	ペットボトル

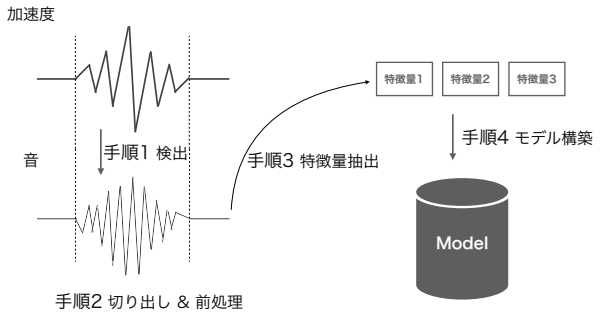

図 2 モデル構築の設計

3.2 特徴量抽出

音から抽出する特徴量の一覧を**表 2** に示す．これらの特徴量は，一般的に機械学習で用いられている特徴量であるため，本システムで採用した．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である．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である．Melspectrogram 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である．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 [12]．

表 2 特徴量一覧

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

4. 評価実験

4.1 実験目的

ユーザがごみを叩く動作をするだけで，ごみの種別・位置を記録するシステムを設計するうえで，叩いた音からごみの種別を推定する精度が極めて重要であると考える．ごみの種別の推定が誤認識していた場合，ユーザは，自身でラベリングの設定を変更するため手間がかかってしまい，研究の目的を満たすシステムを構築することが困難になる．有効な手法は，各ユーザに対して個別のモデルを構築することだが，それには大きなコストがかかる．もし，一定数の学習データを用いて構築したモデルで，未学習のシステム利用者のごみを叩いた音の種別推定の精度が高い場合，各ユーザに個別のモデルを構築しなくて済む．モデルの精度を向上させる別のアプローチとして，機械学習の手法の選定は重要である．機械学習の手法には，データの特

© 2021 Information Processing Society of Japan

徴によって分類が得意・不得意が存在するため，最適な手法を選択する必要がある．そして，システムの運用を考える上で，環境音が入った実空間でもモデルの推定精度が高いことは不可欠である．ごみを拾う人がごみを叩くとき，様々な種類の環境音が入り込む可能性が高く，それらの音に対応できるようなモデルを構築する必要がある．

そこで本実験では，システム利用者のごみを叩いた音からごみの種別を推定するタスクにおいて，各ユーザのごみ拾いした際の音データのみから学習を行った推定器を用いる方式（以降，InSub 方式）と，他のユーザのごみ拾いした際の音データから学習を行った推定器を用いる方式（以降，LOPO 方式）の 2 方式において，どの程度違いがあるのか検証する．同様に，どの機械学習の手法が最も精度が高いかを検証する．また，静寂な空間で構築したモデルと，環境音が入った音で構築したモデルにおいて，どの程度の差が生じるかを検証する．

評価には Accuracy，Precision，Recall，F 値を用いる．Accuracy（正解率）は，ごみの音から種別を推定したデータのうち，システムによって正しく推定されたデータの割合である．Precision（適合率）は，ごみの音から種別を推定したデータのうち，実際にその種別のごみであったデータの割合である．Recall（再現率）は実際にごみの音から種別を推定したデータのうち，システムによって正しく推定されたデータの割合である．F 値（F-measure）は，Precision と Recall の調和平均であり，次式で表される．

$$F = 2 \times \frac{Recall * Precision}{Recall + Precision}$$

4.2 データセット

4.2.1 静寂な環境で収集したデータセット

以降，データ収集に協力してくれた実験参加者のことを計測者と称し，計測者が収集した 1 つの音データをサンプルと称する．計測者がごみを叩く音は，スマートウォッチに内蔵されているマイクを用いて計測した．今回実験で使用したスマートウォッチは，Apple watch series 5 である．データ収集時のサンプリングレートは，44100Hz とした．実験者は，表 1 の 5 種類のごみを，10 個ずつ用意した．計測者は，専用のデータ収集アプリ（**図 3**）でサンプルを収集する．計測者は 20 代の大学生 12 名（男性 11 名，女性 1 名）である．実験者は，収集した音データから，専用のソフトウェアを用いて推定モデルを構築した．計測者は，静かな環境でデータ収集を行った．下記の手順で計測を行い，1 人 250 件，全体で 3000 件のサンプルを収集した．

Step1 本システムを利用しているシーンに近づけるため，計測者は，はじめに軍手を装着する．また，音の情報を収集しやすくするために，本実験では初期検討として，スマートウォッチを計測者の利き腕（ごみを叩く方の腕）に装着した．

図 3 データ収集アプリの画面

図 4 実験風景

Step2 実験者は，データ収集アプリ（図 3）で計測者の計測者名，ごみの種類をセットする．

Step3 計測者は，データ収集アプリ（図 3）の「データ収集」ボタンを押下し，ごみを 3 回を叩く．

Step4 計測者は，Step3 の動作を，1 つのごみにつき 5 回繰り返す（**図 4**）．

Step5 計測者は，1 つのごみの種類を 10 個全て叩いたら，Step2 に戻る．そして実験者は，ごみの種類を再び別の種類に設定して，以降 Step3〜Step4 を同様に繰り返す．計測者が全てのごみの種類を叩いたら，「データ転送」ボタンを押下してサーバーに送信し，実験は終了となる．

4.2.2 環境音を合成したデータセット

収集した 3000 件のサンプルを環境音と合成した音を作成し，モデルの評価を行う．環境音は，フリーサイトから 100 個のデータをダウンロードして利用した [13]．専用のソフトウェアを用いて，環境音のデータ 100 個からランダムで 5 個を選択し，1 件のサンプルと 5 つの環境音それぞれで合成音を作成する．上記の操作を全てのサンプルに行うことにより，計 15000 件のサンプルを作成した．

© 2021 Information Processing Society of Japan

4.3 機械学習の手法

本実験で利用する機械学習の手法の一覧を**表 3** に示す．Random forest は，決定木を大量に生成し，各決定木の結果を集計して予測する，アンサンブル学習の手法の 1 つである．Random forest は，汎化性能が高く，処理の並列性も高いのが特徴である．MLP は，入力層と隠れ層と出力層が全結合である，もっとも単純なディープラーニングの手法の 1 つである．Logistic regression は，線形回帰の出力をロジット関数に入力することで，分類問題に対応した回帰モデルである．決定木は，分類木と回帰木を組み合わせたもので，ツリー（樹形図）によってデータを分析する手法である．Extra tree は，基本的には Random forest と同じアルゴリズムであるが，木のノードを分割する際に，特徴量をランダムに選ぶことが特徴である．ナイーブベイズは，ベイズの定理を利用した高速な分類アルゴリズムで，高次元のデータセットに適している．特に，Gaussian Naive Bayes は，連続データに利用する分類器である．SVM は，マージン最大化を取り入れることで，少ないデータでも汎化性能が高い分類回帰モデルを実現できる手法である．KNN は，説明変数からデータ間の距離を計算し，分類するデータに近い k 個の分類内容から，多数決で属するクラスを推定する．LightGBM は，決定木アルゴリズムに基づいた勾配ブースティングのフレームワークである．推論速度が早いという特徴がある．

表 3 機械学習の手法

手法名	特徴
Random forest	決定木，バギング
MLP	ニューラルネットワーク
Logistic regression	回帰モデル
Decision tree	決定木
Extra tree	決定木，バギング
Gaussian Naive Bayes	分類モデル
SVM	回帰モデル
KNN	分類モデル
LightGBM	決定木，勾配ブースティング

4.4 実験方式

4.4.1 InSub 方式（w/o noise）

計測者 1 人あたりの教師ありサンプル 250 件に対し，前処理として最大値 1，最小値 0 の正規化を行う（手順 2）．これにより，ごみを叩く際に生じる計測者ごとの振幅の差を無くすことができる．前処理した音データから，特徴量（251 次元）を算出する（手順 3）．抽出した特徴量を用いてモデルの構築を行う（手順 4）．構築したモデルの評価には 10 分割交差検定法を用いる（Leave-One-Garbage-Out cross validation）．機械学習の手法は，表 3 の手法全てを用いて検証する．具体的には，250 件の教師ありサンプルを 10 分割し，225 件を学習サンプル，25 件を推定サンプルに分けて Accuracy・Precision・Recall・F 値を算出する．

4.4.2 LOPO 方式（w/o noise）

計測者以外のサンプル 2750 件全てをマージして，計測者 250 件のサンプルからごみの種別を推定するタスクを行う（Leave-One-Person-Out cross validation）．上記以外の処理は，InSub 方式（w/o noise）と同様である．

4.4.3 InSub 方式（w/ noise）

4.2.2 項で作成したデータセットを用いてモデルを構築する．具体的には，1250 件のサンプルを 10 分割し，1125 件を学習サンプル，125 件を推定サンプルに分けて Accuracy・Precision・Recall・F 値を算出する．上記以外の処理は，InSub 方式（w/o noise）と同様である．

4.4.4 LOPO 方式（w/ noise）

4.2.2 項で作成したデータセットを用いてモデルを構築する．具体的には，計測者以外のサンプル 13750 件全てマージして，計測者 1250 件のサンプルからごみの種別を推定するタスクを行う．上記以外の処理は，InSub 方式（w/o noise）と同様である．

4.5 結果

実験結果を**表 4**，**表 5**，**図 5**〜**図 7** に示す．表 4，表 5 に注目すると，Random forest，MLP，LightGBM の F 値の精度は，どの実験方式においても上位 3 位であることから，3 手法のうちどれかが最適な機械学習の手法であることを判断することができる．したがって，Random forest，MLP，LightGBM の精度（F 値）に対して検定を行う．InSub（w/o noise）方式の手法間において，一元配置分散分析を行ったところ，10 %水準で有意傾向があった．各手法間の結果に対し，Tukey の多重比較比較検定を行ったところ，手法間で差はみられなかった．InSub（w/ noise）方式の手法間において，一元配置分散分析を行ったところ，1 %水準で有意差がみられた．各手法間の結果に対し，Tukey の多重比較比較検定を行ったところ，LightGBM と MLP，Random forest と MLP の 2 手法間において 1 %水準の有意差がみられた．また，LightGBM と Random forest の 2 手法間において 10 %水準の有意傾向があった．LOPO（w/o noise）方式の手法間において，一元配置分散分析を行ったところ，10 %水準で有意傾向がみられた．各手法間の結果に対し，Tukey の多重比較比較検定を行ったところ，LightGBM と MLP，の 2 手法間において 10 %水準の有意傾向がみられた．LOPO（w/ noise）方式の手法間において，一元配置分散分析を行ったところ，5 %水準で有意差がみられた．各手法間の結果に対し，Tukey の多重比較比較検定を行ったところ，LightGBM と MLP の 2 手法間において 5 %水準の有意差がみられた．また Random forest と MLP の 2 手法間において 10 %水準の有意傾向があった．

© 2021 Information Processing Society of Japan

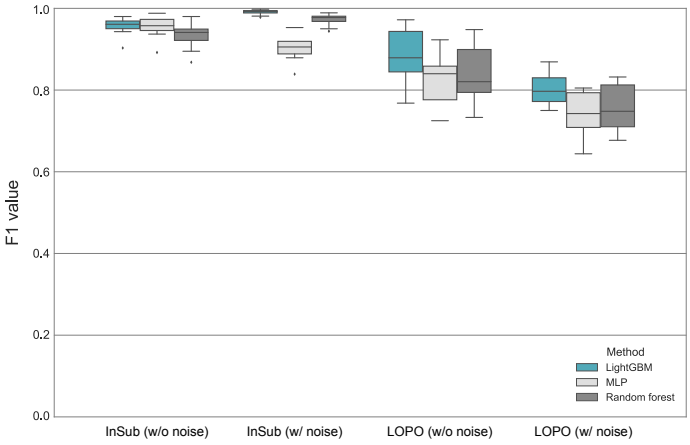

図 5　モデルの評価（F 値）

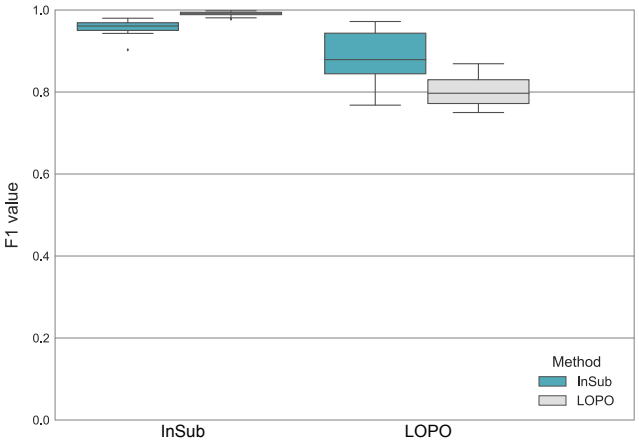

図 7　環境音の有無における評価（F 値）

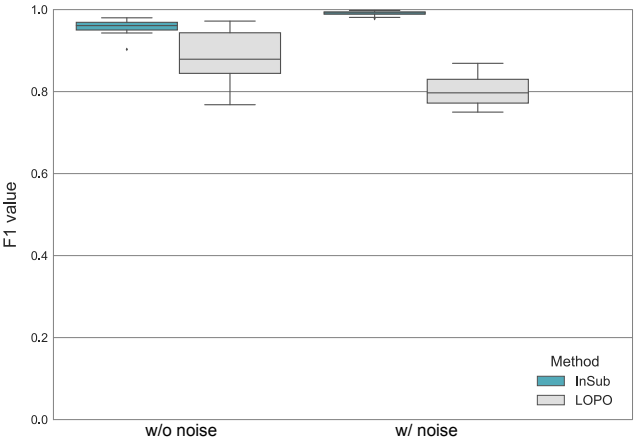

図 6　InSub 方式，LOPO 方式間における評価（F 値）

以降，InSub 方式，LOPO 方式間における評価と，環境音の有無における評価において，機械学習の手法は，検定結果から最も有意な手法であった LightGBM を採用した．はじめに，方式間における精度の結果に注目する．方式間の結果（**図 6**）に対し，対応のある t 検定を行ったところ，InSub（w/o noise）方式と LOPO（w/o noise）方式において，InSub（w/o noise）方式が，5 %水準で有意に上回っていた．また，InSub（w/ noise）方式と LOPO（w/ noise）方式において，InSub（w/ noise）方式が，1 %水準で有意に上回っていた．つぎに，環境音の有無における違いの結果（図 7）に対して対応のある検定を用いたところ，InSub（w/o noise）と InSub（w/ noise）方式において，InSub（w/ noise）方式が，1 %水準で有意に上回っていた．また，LOPO（w/o noise）方式と LOPO（w/ noise）方式において，LOPO（w/o noise）方式が 1 %水準で有意に上回っていた．

4.6　考察

表 4，表 5，と検定の結果から，LightGBM が最も有意なモデルであることが判断できる．また，同様に Random forest の精度も高い．これより，決定木を用いたアンサンブル学習の手法が，叩いたごみの音から種別を推定するタスクにおいて高精度で分類できるという知見を得られた．

環境音が入ってないデータセット（w/o noise）と環境音が入っているデータセット（w/ noise）におけるモデルの評価から（図 7），InSub 方式は，環境音が入っているデータセットで構築したモデル（w/ noise）の方が，環境音が入っていないデータセットで構築したモデル（w/o noise）よりも高い精度が出ることがわかった．これより，ユーザのデータのみで構築したモデルのデータの数を増やすことで，環境音の有無にかかわらず，モデルの精度が向上する可能性が示唆された．

InSub 方式と LOPO 方式の手法間（図 6）では，環境音の有無にかかわらず，InSub 方式の方が LOPO 方式よりもモデルの精度が高いことが検定結果からわかる．ごみを叩く音は，計測者によって叩き方は違うため，方式間における精度の差は想定通りである．しかし，LOPO 方式（w/o noise）では精度（F 値）が 88.5 %，LOPO 方式（w/ noise）では精度（F 値）が 80.2 %出ているため，決して低くはない．また，3.1 節でも述べたように，システムが推定したごみの種別の確信度の低い場合，システムがユーザ自身にごみの種別を聞くことで，誤認識を防ぎ，ユーザの叩き方にモデルが適用していく設計にすることで問題を解決できる可能性が高い．

最後に，システムの分類精度をより向上させるために，LOPO 方式（w/ noise）における各計測者の傾向を考察する．計測者の傾向は大きく 2 つに分かれた．1 つ目の傾向は，計測者 A のような結果で（**図 8**）高精度にごみの叩く音から種別を推定することが可能であるパターンである．2 つ目の傾向は，プラスチックを紙類と誤認識してしまい，全体の精度が低くなる結果である（**図 9**）．缶や瓶，ペットボトルは素材が統一されているためモデルの精度は高くなる．一方で，紙類は，紙コップやティッシュ，プラスチックはお菓子袋やビニール袋など素材が異なるごみでモデルを構築しているため，ユーザが叩く音の微妙な違いが生じ，

© 2021 Information Processing Society of Japan

表 4　InSub 方式

Method	w/o noise				w/ noise			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F
Random forest	0.936	0.946	0.936	0.934	0.972	0.974	0.973	0.973
MLP	0.957	0.964	0.958	0.957	0.904	0.908	0.904	0.903
Logistic regression	0.922	0.933	0.922	0.920	0.922	0.933	0.922	0.920
Decision tree	0.812	0.833	0.812	0.801	0.786	0.791	0.785	0.785
Extra tree	0.666	0.690	0.666	0.666	0.790	0.794	0.790	0.789
Gaussian Naive Bayes	0.836	0.853	0.836	0.823	0.706	0.730	0.706	0.699
SVM	0.531	0.533	0.531	0.481	0.534	0.586	0.533	0.523
KNN	0.771	0.791	0.771	0.762	0.578	0.586	0.578	0.570
LightGBM	0.958	0.965	0.958	0.957	0.958	0.965	0.958	0.957

表 5　LOPO 方式

Method	w/o noise				w/ noise			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F
Random forest	0.846	0.859	0.846	0.842	0.76	0.772	0.761	0.754
MLP	0.84	0.862	0.84	0.837	0.751	0.773	0.751	0.747
Logistic regression	0.762	0.801	0.762	0.751	0.655	0.68	0.656	0.645
Decision tree	0.69	0.707	0.69	0.686	0.648	0.652	0.648	0.646
Extra tree	0.555	0.571	0.555	0.553	0.507	0.514	0.507	0.505
Gaussian Naive Bayes	0.639	0.693	0.639	0.619	0.578	0.635	0.578	0.56
SVM	0.584	0.629	0.584	0.559	0.565	0.613	0.565	0.56
KNN	0.575	0.603	0.575	0.567	0.488	0.51	0.489	0.485
LightGBM	0.887	0.895	0.887	0.885	0.805	0.821	0.805	0.802

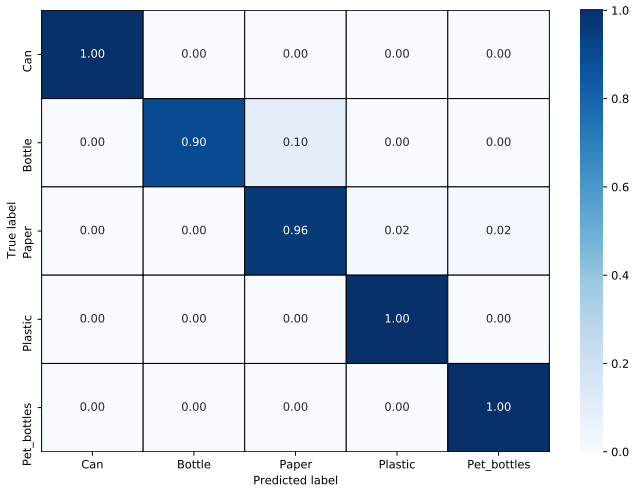

図 8　計測者 A LOPO（w/ noise）

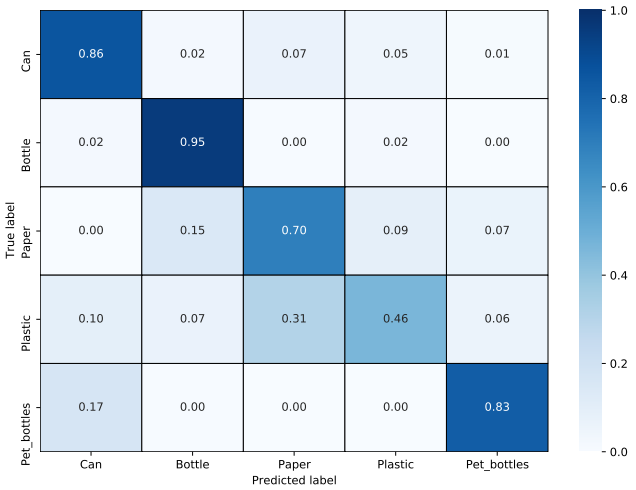

図 9　計測者 B LOPO（w/ noise）

精度が下がってしまう可能性が高いと考える．しかしこの問題は，多くのプラスチックや紙類のデータを収集してモデルを構築することで解決できると考える．

5.　おわりに

本稿では，提案システム [1] のごみの種別を推定するモデルの構築手法を提案した．本実験では，ごみの叩いた音からごみの種別を推定するタスクにおいて，最も精度の高い機械学習の手法，個人のデータで構築したモデルと他者のデータで構築したモデルとの精度検証，環境音の有無によるモデルの精度検証を行った．実験結果より，Random forest や LightGBM などの決定木を用いたアンサンブル学習の手法がごみの音から種別を推定するタスクにおいて，高精度で分類できるという知見を得られた．これより，ユーザのデータのみで構築したモデルのデータの数を増やすことで，環境音の有無にかかわらず，モデルの精度が向上する可能性が示唆された．InSub 方式と LOPO 方式の手法間において，InSub 方式が LOPO 方式よりも有意であるということが明らかになった．実験結果より，精度が低い計測者は，紙類をプラスチックに誤認識してしまってい

© 2021 Information Processing Society of Japan

ることが混同行列から判断できる．しかしこの問題は，多くのプラスチックや紙類のデータを収集してモデルを構築することで解決できると考える．

今後は，表 2 の特徴量だけでなく，音の特徴量を多く利用してごみの種別を推定する．その後に，Weka[14] という最適な特徴量を選択してくれるソフトウェアを利用することによって，最適な特徴量を発見する予定である．また実験結果より，個人のデータのみで構築されたモデルでは，高精度でごみの種別を分類することができたので，実際に屋外でデータを収集しても，高精度でごみを分類することができるのか検証する予定である．その際，計測者は，ピリカ [6] などの手法を利用し，どちらの手法の方が手間がかからなかったかを実証実験していく予定である．また，計測者を増やすことによって，より汎用的なモデル構築に取り組む予定である．

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

参考文献

[1] 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.

[2] Maria Araújo and Monica Costa. A critical review of the issue of cigarette butt pollution in coastal environments. *Environmental Research*, Vol. 172, , 02 2019.

[3] The World Economic Forum. The new plastics economy rethinking the future of plastics. `http://www3.weforum.org/docs/WEF_The_New_Plastics_Economy.pdf`, 2016. Accessed: 2021-05-09.

[4] 井上芳樹, 戸田任重．諏訪湖・天竜川上流における漂着ゴミ．環境科学会誌, Vol. 16, No. 3, pp. 167–178, 2003.

[5] 植田憲, 高野維斗, 神崎広史, 宮崎清．ごみの「ポイ捨て」の未然防止に関する調査・研究:千葉市・駅周辺地域におけるごみ捨て行為の実態調査に基づいて．日本デザイン学会研究発表大会概要集, 第 53 巻, p. 188, 2006.

[6] 株式会社ピリカ．ごみ拾い sns ピリカ—いつでもどこでも気軽にボランティア．`https://sns.pirika.org/`. Accessed: 2020-07-27.

[7] 高橋祐平, 石坂公一, 小地沢将之．タバコのポイ捨てポテンシャルの分布構造－仙台市中心部のアーケード街を対象として－．日本建築学会技術報告集, 第 15 巻, pp. 257–260, 2009.

[8] 早瀬光司, 錫木圭一郎．公共空間における散乱ごみ防止のための, 実験系の設定, 散乱ごみの分析, 人員の動態計測およびごみ発生原単位に関する基礎的研究．廃棄物学会論文誌, Vol. 9, No. 6, pp. 274–280, 1998.

[9] 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.

[10] Taesik Gong, Hyunsung Cho, Bowon Lee, and Sung-Ju Lee. Knocker: Vibroacoustic-based object recognition with smartphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Vol. 3, No. 3, 2019.

[11] 食品容器環境美化協会．散乱実態調査（2016 年度）—公益社団法人食品容器環境美化協会．`https://www.kankyobika.or.jp/recycle/research/3R-2016`, 2016. Accessed: 2020-07-27.

[12] 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.

[13] 株式会社 VSQ（ブイエスキュー）. 効果音フリー素材　環境 — special — vsq. `https://vsq.co.jp/special/se_environment/`. Accessed: 2021-04-27.

[14] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, Vol. 11, No. 1, p. 10–18, November 2009.

© 2021 Information Processing Society of Japan