

# 屋外移動支援を目的とした国内特化型データセット VIDVIPの特徴と運用設計

馬場 哲晃<sup>1,a)</sup>

**概要:** 本稿では、2018年より開発を継続してきた、視覚障害者の屋外移動支援を目的とした物体検出データセットに関する開発の経過と一般公開の概要を報告する。さらに収集したデータセットの特徴を俯瞰し、開発の過程から明らかになった、国内における地域毎の歩道整備状況や問題点を整理し、今後の運用設計や一部実装したアプリケーション等について報告する。

## 1. 背景

本研究では視覚障害を持つ当事者が、単独で屋外移動を行うためには今後深層学習を用いたコンピュータビジョン技術が極めて重要であると考え、そのための基礎となるデータセットの開発に取り組んでいる。国内特化型の物体検出データセット開発を2018年4月より進めており、2021年3月までは物体検出用途、2021年4月からは領域分割用途のアノテーションを行っている。物体検出用途は現時点で32,036枚の画像に対して538,747インスタンスの登録を完了した。実際に登録された画像例を図1に示す。なお、2019年12月までの開発経過（当時の登録画像数7,122枚、インスタンス数は83,319である）及び、それらを利用したナビゲーションシステムによるユーザ実験に関しては文献[1][2]を参照されたい。我々の開発したデータセットはCOCOにおける250万インスタンス[3]やOpen Images Dataset[4]の900万インスタンス数と比較すると小規模ではあるが、PASCAL VOCにおける27,450インスタンスと比較すると検出ネットワークの学習に十分な数量と推察される。

## 2. 関連研究

本研究の実施において、歩道上を目的地に向かって障害物などを避けつつ移動する観点から、自動運転用途のデータセットが本研究が目的とする形式に類似していると考えられる。研究開始当初の2018年代では自動運転用途のオープンデータセットはKITTI [5]がよく知られていたが、2020

年代に入り大手企業からWaymo [6], DriveSeg [7], A2D2 [8]等が公開されている。KITTIでは車や歩行者、バイク等の9クラスが登録されているほか、Waymoでは車、歩行者、標識、自転車搭乗者の4クラス、DriveSegでは電柱や植生等を含めた12クラス、A2D2では3D物体検出ではあるが、建物や信号機、横断歩道や歩道等を含めた38クラスが登録されている等、登録されているクラスラベルはデータセットにより異なる。これら既存データセットを活用することで本研究の目的とする検出器を一定のレベルで構築することが可能であると考えられるが、撮影画像が歩行者目線ではなく自動車運転者目線であり、障害当事者が必要とする情報（段差や歩行者用押し釘、点字ブロック等）がそもそも登録されていない。

本研究で対象とするデータセットは視覚障害者の屋外移動（主に歩道上）に必要な検出物を対象としているが、車椅子では自動運転、近い未来ではヘッドセットや透過型メガネディスプレイ等利用時における安全センシング用途等に応用が可能である。Kaurら[9]はウェブカメラに測距センサを追加し、RGB+Depthの1チャンネルを追加したFaster R-CNNベースのネットワークを設計し、周辺情報（人や車）等を音声ナビゲーションを通じてフィードバックを行った。Mulfari [10]らはシングルボードPC(Raspberry Pi 3)にtensorflow環境を構築し、メガネに搭載したカメラから物体検出を行うシステム実装を行い、視覚障害者支援に関する可能性を議論している。Chaudhryら[11]は顔認識を利用した視覚障害者のための人物特定支援システムを開発している。

歩行者や車椅子利用者が屋外で移動する際の目線撮影によるデータセットはほとんど公開されないだけでなく、学習済モデルも公開されているものは極めて少ない。また

<sup>1</sup> 馬場哲晃  
Tokyo Metropolitan University, 6-6 Asahigaoka, Hino, Tokyo  
191-0065, Japan  
<sup>a)</sup> baba@tmu.ac.jp

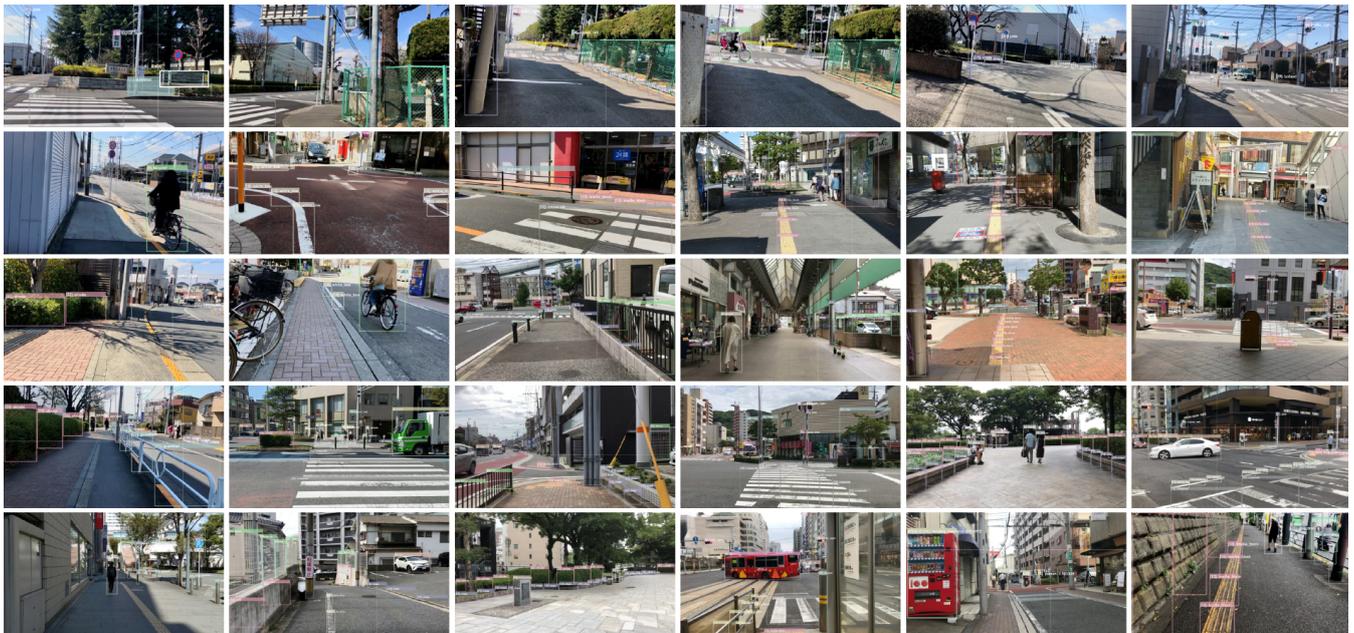


図 1: 実際にアノテーションした各画像の例

公開されていた場合であっても実際の利用状況を想定した画像群で無い等の問題がある。前者は法律による問題も考えられるが、著作権法 47 条に基づく限りは、少なくとも後者については国内法においては配布可能であると考えられる [12]. Thakurdesai[13] らが汎用データセットである COCO と KITTI を利用して屋外歩行者向けの物体検出器を開発したことは先に述べたが、国内では、東らは自動運転車椅子 [14] の自立移動の経路計算に GAN を利用しているが、データセットを独自に開発している。大北らも独自にデータセットを開発し、視覚障害者 [15] ナビゲーションシステムの実装を行っている。これらはデータセット開発が目的ではない為、データセットが公開されないばかりか、局所的な学習モデルであるために汎用性が低く、今後同様の開発を実施する研究者にとって、データセット開発という観点からメリット享受は望めない。

以上のことから本研究ではデータセットの開発に主眼を置き、学習済みモデル及びデータセットを公開することを具体的な研究のゴールとしている。特に本稿ではデータセットの質を検証することで、他の汎用データセットとの違いや、国内の地域ごとの特性に関してまとめる。

### 3. データセットの比較

本データセットと比較的規模が類似している COCO や PASCAL VOC, 自動運転用途の観点から KITTI の 3 つのオープンデータセットと今回開発したデータセット, VIDVIP に関して比較を行った。一般的にデータセット品質や性能には量が重要であるが、その他バウンディングボックスの精度等の指標がある。Li ら [16] らはリモートセンシングデータセットのベンチマークにおいて、データセット

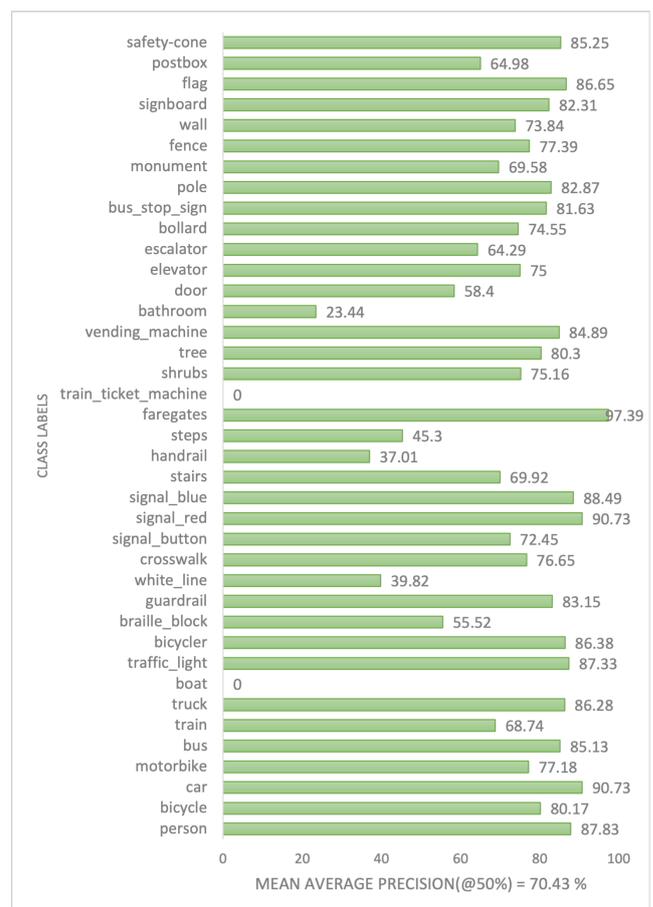


図 2: YOLOv3 による精度一覧

品質の指標の一つとしてインスタンス数との他、mAP や boxplot (箱ひげ図) を参照していることから、本研究においてもそれに倣って議論を進める。各データセットの概要一覧を表 1 に示す。本データセットとクラスラベルがそ

それぞれ VOC, COCO, KITTI のラベルが共通として考えられる一覧を表 2 に示す。ただし KITTI に関しては Van は Car, Pedestrian 及び Person\_sitting は person, Cyclist は bicycler, Tram を bus として読み替えている。なお本稿では 2D Object Detection に限った比較である。

### 3.1 VIDVIP データセットの概要

VIDVIP データセットは 2 次元物体検出 (Object Detection) と領域分割 (Semantic Segmentation) の 2 種類を含む。領域分割は現在作業中のデータセットであり、本稿では前者の 2 次元物体検出部分に関して主にまとめる。登録クラスラベルは 39 種類あり、これらのラベル選定に関しては文献 [1] を参照されたい。図 3 上下にそれぞれクラスラベルごとの登録インスタンス数及び箱ひげ図による登録面積分布図を示す。登録面積分布図の縦軸は登録されたバウンディングボックスにおけるその画像サイズとの比率のことを差しており、1.0 に近いほど画像と同じサイズでアノテーションが行われている。物体検出には主に yolov3[17] を利用して精度確認を行っている。図 2 に各クラスラベルの average precision と全体の mAP を示す。なお mAP の算出にはデータセット全体の 10% を test データに用いている。

アノテーションの品質は重要であるが、認識ラベルの誤りが無いことはもちろん、精確なアノテーションがなければ障害物のズレが生じたり、安全性の高いガイドを提供することが困難になる。大規模データセット開発にはしばしばクラウドソーシングが活用されるが質の低下が近年問題視されている。例えば画像データセットとして有名な ImageNet では全体の 6% にあたる 2916 箇所のラベル誤りが報告されている [18]。これまでのアノテーション作業においては、研究グループ内において、著者研究室に所属する本業務に熟練した学生が継続的に実施してきた。さらには研究室内で開発したアノテーションアプリケーションを利用し、一枚一枚をすべて 2 重工程による目視確認を行っている。

国内における歩道移動支援を目的としているため、様々な地域からの画像収集が望ましいが、地理的、実験環境的な制約から大きく分けて、東京都 (日野市, 武蔵野市, 八王子市, 文京区, 北区), 神奈川県 (海老名市, 平塚市), 愛知県 (北名古屋, 岩倉市, 稲沢市, 名古屋市, 蟹江町, 豊川市), 広島平和記念公園 (広島県広島市), 福岡県北九州市の画像郡を中心としたデータセットになっている。それぞれの地域ごとや場所における傾向については後述する。

### 3.2 比較

PASCAL VOC は比較的小規模なベンチマークデータセットとして広く知られている。今回比較に利用したものは Pascal VOC Challenges の中で提供されていた中で最

表 1: 比較するデータセットに関する概要一覧

Datasets	Categories	Images	Instances	Year
VOC	20	17,125	23,979	2012
COCO	80	82,081	255,403	2014
KITTI	9	7,481	38,503	2012
VIDVIP(ours)	39	32,036	538,747	2021

表 2: 共通するクラスラベル比較。

Labels	VIDVIP(Ours)	VOC	COCO	KITTI
person	○	○	○	○
bicycle	○	○	○	○
car	○	○	○	○
motorbike	○	○	○	
bus	○		○	○
train	○	○	○	○
truck	○		○	○
bicycler	○			○
traffic light	○		○	

もデータ量が多い The VOC2012 Challenge によるものを利用した。このデータセットでは認識と物体検出に関するデータセットとなっており、20 クラスラベルに関して、表 1 に示す通り、17,125 枚の画像、23,979 インスタンスのアノテーションとなっている。図 4 左は、共通する各クラスラベルに関するバウンディングボックス面積の箱ひげ図である。1.0 になるほど登録された物体が画像サイズと等しいことになる。いずれのクラスも比較的大きく撮影されているが、car に関しては他のインスタンスに比べて面積比率が小さい。図 4 右は各クラスラベルのインスタンス数となっており person が他と比較して多く登録されていることがわかる。なお、truck や bicycler が 0 であるが、元々 VOC データセットにこれらのクラスラベルは含まれないためである。

COCO は Microsoft 社により提供されているベンチマークデータセットである。今回比較に利用したものは 2014 年の学習用データとなっている。このデータセットは 2021 年 5 月 5 日現在 2017 年で主たる更新は止まっており、2017 年段階では 12 万枚程度まで画像データが増加しているが、今回は検討の扱いやすさから 2014 のデータを利用した。ただしアルゴリズムコンペティションである Object Detection Task は 2020 年においても開催されており、以前として活発なコミュニティであることが伺える。図 5 に面積比率及びインスタンス数を示す。いずれも VOC データセットと似た分布になっていることがわかる。

KITTI はこれまでの VOC や COCO と趣旨が異なり、自動運転用途のデータセットとなっている。走行時の車内からの視点のみで撮影された画像に対してアノテーションがされている。ステレオ画像や Laser Points of Clouds によるデータセット等、複数センシングを利用したデータが

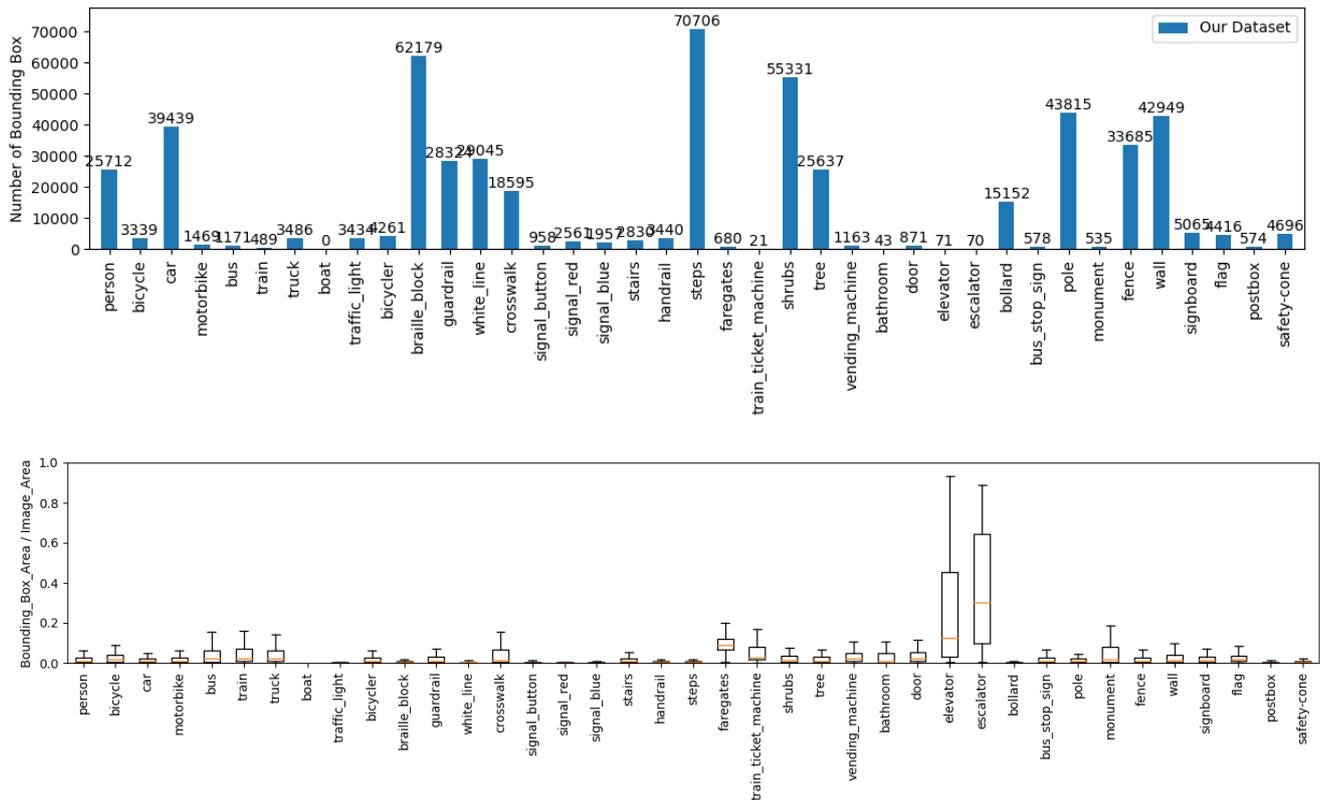


図 3: VIDVIP データセットのクラスラベルごとの登録インスタンス数 (上) と箱ひげ図による面積分布 (下)

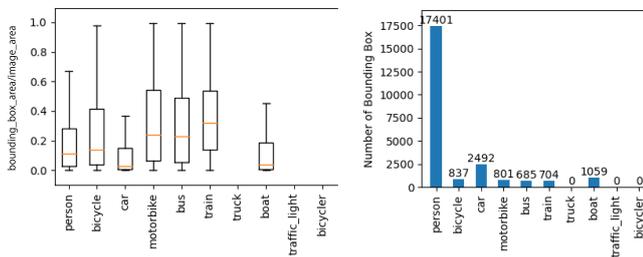


図 4: VOC データセットのクラス毎におけるバウンディングボックス面積の箱ひげ図, インスタンス数

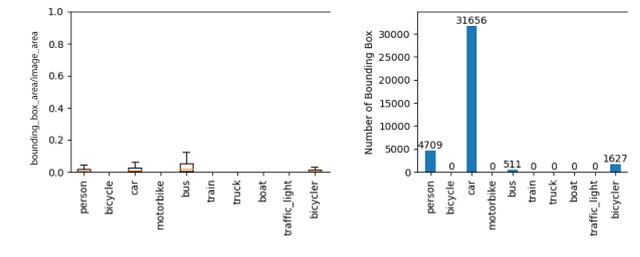


図 6: KITTI データセット (2012) のクラス毎におけるバウンディングボックス面積の箱ひげ図, インスタンス数

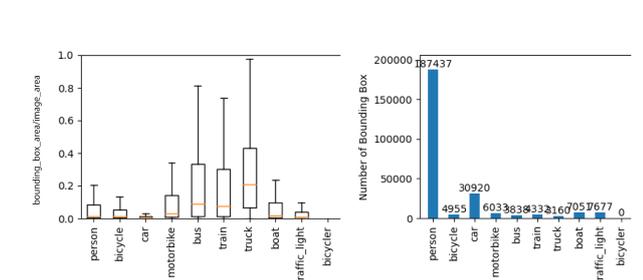


図 5: COCO データセット (2014) のクラス毎におけるバウンディングボックス面積の箱ひげ図, インスタンス数

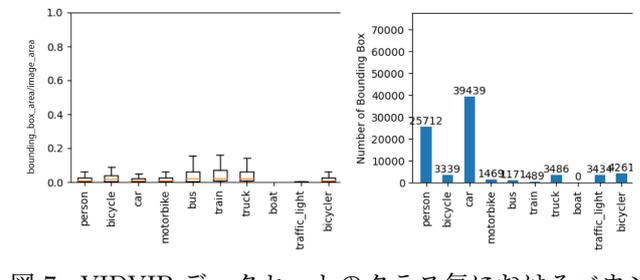


図 7: VIDVIP データセットのクラス毎におけるバウンディングボックス面積の箱ひげ図, インスタンス数

公開されていることに特徴がある。今回はこの中の 2D 物体検出データセットを利用した。図 6 は面積比率及びインスタンス数となる。VOC や COCO と比較して大きく異なる点は面積比率の小ささである。走行中の車内からのカメラ

ラ画像であるため、画面いっぱいに検出対象物体が映ることがない為であると考えられる。インスタンス数に関しては car が最も多く、ついで person, bicycler, bus となっている。インスタンス数が 0 なのは対応するラベル付がないためである。

以上のことから我々のデータセット VIDVIP では、移動中の歩行者目線画像であることから KITTI のように VOC や COCO と比較して面積比率が小さくなることが予想される。図 7 にこれまで VIDVIP データセットの面積比率及びインスタンス数を示す。KITTI データセットと比較するとやや面積比率が大きいことがわかる。これは歩行時の撮影映像と自動車走行時の撮影とでは検出物体対象物と撮影者の距離が異なるためであると推察される。またインスタンス数の分布も KITTI に類似している。これらのことから VIDVIP は自動運転用途のデータセットに近い性質を持ち、検出対象物が車からの撮影と比較して撮影者と距離が近いことが示された。

#### 4. 地域・場所

データセット作成において、特に画像認識においてはその地域ごとの差についても十分に配慮しなければならない。とりわけ本研究は視覚障害者の屋外移動支援であることから、データセットの品質には十分注意をして開発を進めていく必要がある。今後も継続してデータセット開発を行っていきにあたり、地域や場所における特性を把握することで登録されたインスタンスの妥当性を判断することが可能になるだけでなく、今後撮影が必要な場所選定に関しても地域・場所特性は判断材料として利用ができる。そこで本節では VIDVIP データセットにおける各地域性や、場所性に関して議論する。図 8 に各撮影地域毎のアノテーション結果（一部抜粋）をまとめる。愛知県のみ市区町村表記がないのは、複数の区市町村画像がバラバラに保存されたためであり、あくまで参考データとして併記している。横軸は地域、縦軸は全体インスタンス数における任意のクラスラベルインスタンスが占める割合である。例えば東京都武蔵野市の person インスタンスは武蔵野市全体のデータセットにおいて、10%を占める。person に関しては、武蔵野市、広島平和記念公園、愛知県の3つが突出して10%以上の値となっている。人口密度で考えると、東京都文京区も同様に高い数値を示すはずであるが、5%弱である。これは撮影した地域だけでなく、撮影場所に大きく依存していると考えられる。例えば広島平和記念公園は観光地であり、撮影当日も歩行者が多かったためである。また、東京都武蔵野市では主に吉祥寺市街地を撮影したものが多く、愛知県については駅構内での撮影画像が含まれているなどの状況が確認できた。person の増減は撮影場所に大きく依存することがわかる。同様に car では広島平和記念公園が極めて少ない割合であるが、これは公園内での撮影画像が多いためである。

braille\_block（点字ブロック）に関しては、神奈川県海老名市の割合が極めて少ないが、これは撮影画像がほとんど住宅街であったためである。また広島平和記念公園も他の場所と比較して点字ブロックの登録割合が低い。一般的

に公園等は点字ブロック舗装が完備されているような印象があるが、我々の調査からも広島平和記念公園の点字ブロックは原爆ドーム敷地内には設置されておらず、広島平和記念公園内も一部の設置にとどまっていたことがわかっている [19]。一方で市街地の歩道部分の撮影画像が多い地域では15%以上の登録割合であることが確認できる。

段差を示す step では広島平和記念公園の登録割合が著しく高いことが確認できる。これは公園内の植樹がすべて段差ブロックによって区切られているためである。いずれの場所でも屋外を移動することにおいては一定の段差があることが確認できる。

車止めを示す bollard は神奈川県平塚市が高い割合を示している。これは撮影された画像は市街地の中でも比較的車道と歩道の境目がガードレール等で区切られていない場所であり、代わりに車止めが多用されていることが原因と考えられる。guardrail（ガードレール）の値を見ると、同様に平塚市の登録割合が小さいことがわかる。fence や wall は神奈川県海老名市での撮影データにおいて突出して登録割合が大きい。これは歩道のない住宅街の画像であるためと考えられる。自宅の塀や郊外の畑における fence 等が多く撮影されている。

signboard は店の軒先にある置型看板であり、視覚障害の当事者が通行する際に障害物となるために登録をしている。同様の理由でのぼり旗を flag として登録している。神奈川県平塚市と東京都文京区にて一定の登録割合を示している。東京都文京区の画像は、JR 御茶ノ水駅から東京大学本郷キャンパスに向かうまでの経路が収録されている。いずれも市街地及び商店街の歩行であるが、文京区では置き看板、平塚市ではのぼり旗が利用される傾向があることがわかる。

開発したデータセットを撮影場所毎に比較をすることで、公園や市街地、住宅街、歩道経路等の違いを確認することができた。データセットの品質確認において有益な情報であるのは明らかであるが、物体検出器と組み合わせることで、指定ルートを歩行するだけで、点字ブロックやガードレールの出現頻度と登録データセットの統計情報から公共整備に役立てるなどの応用も考えられる。

#### 5. 多様なデータの収集

本研究では研究室内及び共同研究先と連携して5-6人程度の小規模グループでアノテーション作業を行ってきた。一方で、様々な地域の特徴を網羅できているとは言い難い。例えば歩行者用押し釦は見慣れた黄色ボックスタイプから、音響用押し釦やタッチ式スイッチ、さらには高齢者用、音響式信号等、微妙に形状や色が異なる。図9はVIDVIPに登録されている歩行者用押し釦画像の一例である。形状や色が異なるだけでなく、軽微な形状、イラストの違い、さらには経年劣化による外観変化等様々なものが観察され

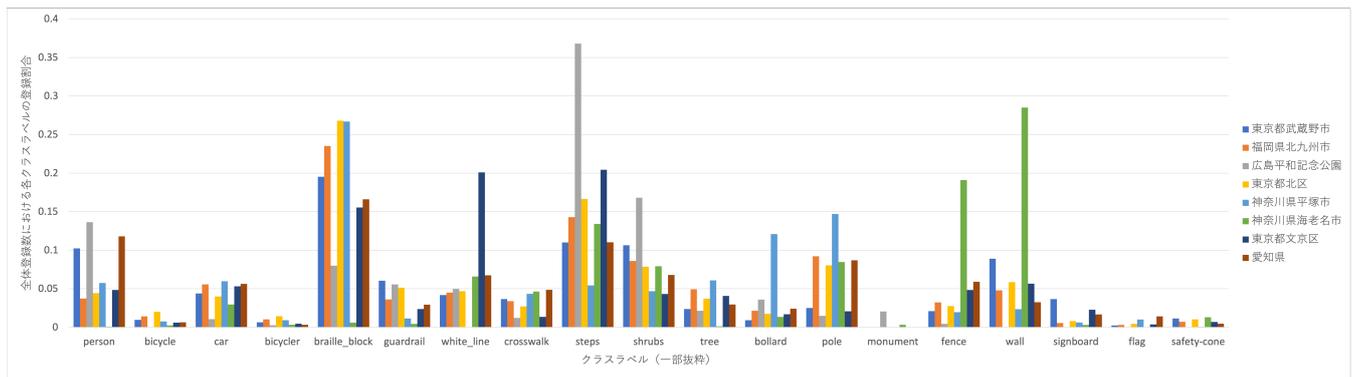


図 8: VIDVIP データセットのクラス毎におけるバウンディングボックス面積の箱ひげ図, インスタンス数



図 9: VIDVIP データセットに登録されている, 各種の歩行者用押し釦

る. この他所謂「ロング押し釦」と呼称される飛び出し式の押し釦の形状は既存のもの\*1である一方で, 深層学習のアルゴリズム上, このようなデータも合わせて登録することが好ましい.

経年劣化した歩行者用押し釦のように, 消えかけている横断歩道, 白線, 破損や脱色した点字ブロック等はデータオーグメンテーションでは容易に生成できないため, より様々な地域における画像収集が課題となる. そこで我々研究グループでは全国各地からのアノテーション用画像を手軽に提供できるシステムとして, スマートフォン用アプリケーションを現在開発している. アプリケーションは iOS 上で動作し, すでにベータ版にてテストしている. 学習済みモデルを利用し, リアルタイムで物体検出結果を確認でき, 認識状況が良くない場合はその場で当該画像を GCP (Google Cloud Platform) 上のクラウドサーバにアップロードすることができる. プロジェクトチーム内で現在テストを行っており, 普段から気になる箇所に関してデータアップロードを繰り返している. スマートフォン上からアノテーションも可能であり, 撮影時の GPS 座標やスマートフォンの撮影方向や姿勢等も合わせて json データとして保

\*1 信号の「ロング押ボタン」を集めました  
<https://nonsensedances.com/?p=4376>  
の記事にいくつかの画像例が掲載されている (最終確認日 2021 年 5 月 5 日)



図 10: 歩道上に点字ブロックが埋設されていない例

存される. 本アプリを学会やイベント等を通じ, 今後は多くの一般ユーザが容易に参画可能な環境を目指している.

多くの画像データを収集する中で, 安全面の観点から懸念のある場所があることがわかってきた. 図 10 は東京都北区の駅周辺で撮影された市街での画像である. この通りは一方通行であるが車が走行可能な道路が右手に, 歩道が左手に確認できる. しかし点字ブロックは道路内に埋設されており, 視覚障害の当事者にとっての場合本当に点字ブロック上を歩くことが安全なのかが不明である. この他同様の安全面での理解が困難な画像が報告されており, このような画像に対してどのようなアノテーションを行う (または行わない) のか, 今後の課題となっている.

## 6. 運用設計

本稿では 3 年間かけて開発したデータセットに関する精度や品質に関して他のベンチマークデータセットとの比較や, 地域性や場所性の観点から検証, 議論を行った. スマートフォンアプリを通じて今後はさらに国内各地の歩行者目線による画像収集を継続していく. すでに学習済みモデルはウェブサイト上にて mlmodel 及び weights 形式にて提供している\*2. ウェブサイト上で公開されている学習済みモデルを利用することで本研究で開発した物体検出器を PC やスマートフォン上で実装ができる. 本モデルを一部

\*2 <https://tetsuakibaba.jp/project/vidvip>

学習データを追加させて、株式会社 CSI は視覚障害者向け屋外ナビゲーションシステムを開発した\*3。すでに100人近くの被験者実験を終えており、今年度中のサービス展開を計画している。このように本開発プロジェクトを通じてサードパーティが容易に視覚障害支援開発に参画可能な環境構築ができた。

以後本研究プロジェクトは学習データセット開発及び、学習済みモデルの公開を主たる目的とし、文献 [20] にて報告したユーザ参加型アノテーションアプリケーションを利用し、国内における様々な画像データ収集を継続して実施する。学習済みモデルを利用したサンプルコード等も今後ウェブサイト上で公開していく。

現在は物体検出だけではなく領域分割によるアノテーション作業も開始しており、車椅子やパーソナルモビリティ等の自動運転や安全制御のためのデータセットや学習モデル提供を行う予定である。今回作成したデータセットは開発チームにより撮影された画像群であり、人物の映り込み等のプライバシー配慮がなされていない為データセットそのものを公開することができない。この問題については現在法律専門家と研究プロジェクト\*4を実施し、スマートフォンアプリケーションによる画像提供に関して利用規約を準備することで解決する予定である。

**謝辞** 本研究は JSPS 科研費 JP18H03486 の助成を受けたものです。

## 参考文献

- [1] Baba, T., Ishisone, K., Watanabe, K., Watanabe, H., Kamae, T., Suematsu, K., Takata, S. and Kuga, Y.: Developing a Localized Object Detection Dataset Supporting Sidewalk Use for Visually Impaired Persons in Japan, *Transactions of the Virtual Reality Society of Japan*, Vol. 25, No. 3, pp. 185–195 (2020).
- [2] Baba, T.: Design for the visually impaired when traveling outdoors using omnidirectional imagery and image recognition, *Impact*, Vol. 2020, No. 7 (2020).
- [3] Lin, T., Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L.: Microsoft COCO: Common Objects in Context, *CoRR*, Vol. abs/1405.0312 (online), available from <http://arxiv.org/abs/1405.0312> (2014).
- [4] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Mallocci, M., Duerig, T. and Ferrari, V.: The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale (2018).
- [5] Geiger, A., Lenz, P., Stiller, C. and Urtasun, R.: Vision meets Robotics: The KITTI Dataset, *International Journal of Robotics Research (IJRR)* (2013).
- [6] Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., Vasudevan, V., Han, W., Ngiam, J., Zhao, H., Timofeev, A., Ettinger, S., Krivokon, M., Gao, A., Joshi, A., Zhang, Y., Shlens, J., Chen, Z. and Anguelov, D.: Scalability in Perception for Autonomous Driving: Waymo Open Dataset, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020).
- [7] Fridman, L. D. J. T. R. S. B. R. L.: MIT DriveSeg (Manual) Dataset (2020).
- [8] Geyer, J., Kassahun, Y., Mahmudi, M., Ricou, X., Durgesh, R., Chung, A. S., Hauswald, L., Pham, V. H., Mühlegg, M., Dorn, S., Fernandez, T., Jänicke, M., Mirashi, S., Savani, C., Sturm, M., Vorobiov, O., Oelker, M., Garreis, S. and Schuberth, P.: A2D2: Audi Autonomous Driving Dataset, (online), available from <https://www.a2d2.audi> (2020).
- [9] Kaur, B. and Bhattacharya, J.: A scene perception system for visually impaired based on object detection and classification using multi-modal DCNN, *CoRR*, Vol. abs/1805.08798 (online), available from <http://arxiv.org/abs/1805.08798> (2018).
- [10] Davide Muldari, A. P. and Fanucci, L.: USING TENSORFLOW TO DESIGN ASSISTIVE TECHNOLOGIES FOR PEOPLE WITH VISUAL IMPAIRMENTS, IADIS International Conference Big Data Analytics, Data Mining and Computational Intelligence 2017 (part of MCCSIS 2017), iadis, pp. 110–116 (2017).
- [11] Chaudhry, S. and Chandra, R.: Design of a Mobile Face Recognition System for Visually Impaired Persons, *ArXiv e-prints* (2015).
- [12] 独立行政法人情報処理推進機構 AI 白書編集委員会: .
- [13] Thakurdesai, N., Tripathi, A., Butani, D. and Sankhe, S.: Vision: A Deep Learning Approach to provide walking assistance to the visually impaired (2019).
- [14] 大輔 東, 雄紀服部, 寛典金子, 基大田中, 雅仁高橋, 央河野, 陽介千田, 洋子大森, 憲一郎堀, 秀雄澁谷, 重明松尾, 靖昭 巽, リチャードリー: 対話型 AI 自動運転車いすを核とした福祉インテリジェントモビリティサービスの開発, 久留米工業大学研究報告, No. 43, pp. 2–12 (2021).
- [15] 拓哉大北, 宏樹兒島, 翔 大井, 睦夫佐野: 深層学習を用いた一人称視点に基づく視覚障害者ナビゲーションシステムの検討, 技術報告 10, 大阪工業大学大学院情報科学研究科, 大阪工業大学大学院情報科学研究科, 立命館大学情報理工学部, 大阪工業大学情報科学部 (2019).
- [16] Li, K., Wan, G., Cheng, G., Meng, L. and Han, J.: Object detection in optical remote sensing images: A survey and a new benchmark, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 159, p. 296–307 (online), DOI: 10.1016/j.isprsjprs.2019.11.023 (2020).
- [17] Redmon, J. and Farhadi, A.: YOLOv3: An Incremental Improvement, *arXiv* (2018).
- [18] Northcutt, C. G., Athalye, A. and Mueller, J.: Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks (2021).
- [19] 渡邊康太, 馬場哲晃, 田村賢哉, 渡邊英徳, 釜江常好: 広島平和記念公園における視覚障がい者の碑巡り支援の基礎検討, 技術報告 2, 首都大学東京, 首都大学東京, 東京大学, 東京大学, 東京大学/スタンフォード大学 (2018).
- [20] 哲晃馬場: ユーザ参加型物体検出データセット構築のためのアプリケーション, 技術報告 4, Tokyo Metropolitan University (2020).

\*3 <https://www.computer-science.co.jp/website/service.php>

\*4 科研費基盤研究 (B), 21H03763, 人工知能 (AI) によるデザイン創作と法的保護-デザイン概念の変化も含めて-