

Ortho-detector: オルソ画像を用いた海岸ゴミ検出手法の提案と 検出タスクの分析

海法 修平^{1,a)} 谷村 朋樹^{2,b)} 陳 寅^{3,c)} 大越 匡^{3,d)} 中澤 仁^{2,e)} 鈴木 もとこ^{4,f)} 山田 暁^{4,g)}

概要：ペットボトル等のプラスチックを含む海岸の漂着ゴミは、有害物質を発生する危険性があるため、回収が求められている。一方で、広い海岸からゴミの種類や量を定量的に把握することが難しいため、船の用意や人員の収集など多大なコストをかけて回収に向かったが、ゴミの量が想定より少なかったということが起こりうる。本研究では、海岸のオルソ画像から物体検出技術を利用して、ゴミの種類ごとに検出することで種類と量を定量的に把握する手法を提案する。提案手法では、ドローンで撮影した画像をオルソ画像に変換した後、深層学習を用いてオルソ画像からゴミの種類ごとに検出する。沖縄の海岸のオルソ画像を利用して精度検証を行い、得られたゴミの種類と量のデータの有用性について議論する。

Ortho-detector: Coastal Garbage Detection Using Ortho-image

SHUEI KAIHO^{1,a)} TOMOKI TANIMURA^{2,b)} CHEN YIN^{3,c)} TADASHI OKOSHI^{3,d)} JIN NAKAZAWA^{2,e)}
MOTOKO SUZUKI^{4,f)} AKIRA YAMADA^{4,g)}

1. はじめに

海岸の漂着ゴミ（漂着ゴミ）は景観を乱すだけでなく、ペットボトル等のプラスチックは有害物質を発生する危険性もあり、早急に回収することが求められる。Jambeckらの調査によると、世界中で年間880万トンのプラスチック

ゴミが海に放流されている [8]。海へ放流されたゴミのうち、プラスチックゴミなどの海水より軽いゴミや中に空気が入っているゴミは、海底に沈むことなく、どこかの海岸に漂着し、それが海岸の漂着ゴミとなる。人が住む島や大陸の海岸に漂着したゴミは、住民の清掃活動により、回収されることが多い。しかし、離島などの住民が少ない、もしくは無人の島では、回収が難しく、放置されている島も多くある。実際に、離島に回収作業に行く際には、ボランティアの収集や人とゴミを運ぶための船の用意などが必要となるため、回収に要するコストは大きい。また、広い海岸の一部の領域に漂着ゴミが密集している場合、効率的に回収作業を行うことが困難であったり、大量のゴミを回収し損ねることがある。そのため、離島での海岸の漂着ゴミの回収作業は、事前にプラスチックなどの有害なゴミの量やその位置を把握した上で、最適なタイミングで、十分な人員と船舶を確保して行う必要がある。

しかし、海岸に漂着しているゴミの量を定量的に把握することは難しい。環境省は、海岸の漂着ゴミに関する統計資料の年度報告を行なっている [14]。この資料には、海

¹ 慶應義塾大学総合政策学部
Faculty of Policy Management Keio University, Fujisawa,
Kanagawa, 252-0882, Japan
² 慶應義塾大学環境情報学部
Faculty of Information and Environment Keio University,
Fujisawa, Kanagawa, 252-0882, Japan
³ 慶應義塾大学大学院政策・メディア研究科
Graduate School of Media and Governance Keio University,
Fujisawa, Kanagawa, 252-0882, Japan
⁴ NTT ドコモ
NTT Docomo
a) kaiho@ht.sfc.keio.ac.jp
b) tanimu@ht.sfc.keio.ac.jp
c) yin@ht.sfc.keio.ac.jp
d) slash@ht.sfc.keio.ac.jp
e) jin@ht.sfc.keio.ac.jp
f) motoko.suzuki.dq@nttdocomo.com
g) yamadaakira@nttdocomo.com

底のゴミなど海岸に漂着しているゴミ以外の情報が掲載されていたり、海のゴミに関しての傾向などの情報がある。しかし、調査対象となっている海岸に限られている点や、調査は基本的に年に一度しか行われてない点、一つの海岸における漂着ゴミの分布を把握できない点から、環境省の資料は離島の回収作業のタイミングを測るという観点においては、有用性が低い。そのため、事前に、広い海岸に対して漂着しているゴミの種類と量と位置を定量的に把握する方法が必要となる。

海岸の漂着ゴミを把握するためには、地上を周回する方法と、上空から撮影する方法が考えられる。地上を人が周回する方法は、調査に多くの労働力が必要となるため、現実的でない。AbdullahらはVirtual Reality (VR)を用いてロボットを遠隔操作することで漂着ゴミの調査と回収を行う方法を提案している [1]。Abdullahらの方法では、調査と回収をロボットに同時に行わせることができる。しかし、海岸には大量には大量の漂着ゴミがあることに加えて、地形の凹凸が激しい場所もあるため、ロボットの回収可能量の限界と安定したロボット作業が困難になることから、実用的でない。上空から撮影する方法では、ドローンを用いることで、上空から海岸全体を容易に撮影することができる。また、オルソ画像と呼ばれる、ドローンの高度を保ったまま上空を移動させて撮影し、撮影された写真を統合した画像を利用することで、小さい漂着ゴミなども識別可能な海岸全体の高解像度画像を作ることができる。オルソ画像とは、写真上の像の位置ずれを無くし、空中写真を地図と同じく、真上から見たような傾きのない、正しい大きさと位置に表示される画像のことである。しかし、これらの高解像度画像から、人手でペットボトルなどの小さな漂着ゴミまで探し出すのは多くの労力を必要とする。

そこで、本研究では、海岸全体を写したオルソ画像から物体検出技術を用いて、自動で漂着ゴミの種類、量、位置を把握するシステム、Ortho-detectorを提案する。本システムでは、まずオルソ画像を小さなパッチ画像に分割する。そして、分割されたパッチ画像に対して、物体検出技術を用いて写っている漂着ゴミの種類ごとに検出する。最後に、全てのパッチ画像を統合して、検出結果を可視化することで、海岸全体の漂着ゴミの分布がわかる海岸ゴミマップを作成する。本システムでは、検出結果を集計することで、漂着ゴミの種類、量、位置を定量的に把握することができ、最終的に作成される海岸ゴミマップによって、それらを定量的に把握することもできる。

物体検出を用いた漂着ゴミの自動検出に伴い、海岸の漂着ゴミの画像データセットを作成した。漂着ゴミの種類として採用したのは、ペットボトル、ブイ、バケツ、縄、網、流木、スプレー缶である。また、ドローンの撮影高度は、5章で検証した高度と精度の関係から、20mとした。

事前実験として、漂着ゴミのデータセットを作成する際

の最適高度の検証を行った。まず、データセット作成のために、撮影高度と精度検証のトレードオフについての実験を行い、20メートルが最適であるという結果が得られた。

事前実験を基にして作成したデータセットに対して、Ortho-detectorを適用し、実際に複数の漂着ゴミを検出する実験を行なった。実験の結果、46%の漂着ゴミが検出できた。また、海岸の漂着ゴミ検出のタスクを分析するために、検出対象をペットボトルのみに絞って、2つの追加実験を行なった。1つ目は、オルソ画像を分割した各パッチ画像に対しての物体検出の際に、何個までの漂着ゴミを許容するかについての検証を行なった。2つ目に、学習した海岸とは異なる海岸に対して検出が可能かどうかの検証を行った。

本稿の構成は以下の通りである。まず、2章では海岸における漂着ゴミの実態と解決の必要性と、現在の漂着ゴミの問題に対する対処方法について述べる。3章では、海岸の漂着ゴミの問題を解決するためのシステムとして、物体検出技術を活用したOrtho-detectorを提案し、その詳細について述べる。4章では、物体検出に利用するためのデータセット作成について述べる。そして5章では、撮影高度と検出精度の関係に関する事前実験について述べる。第6章では、作成したデータセットに対しての提案システムの精度検証について述べ、7章で本論文をまとめる。

2. 漂着ゴミの回収

漂着ゴミは景観を乱すだけでなく、ペットボトル等のプラスチックは有害物質を発生する危険性もあることから、早急な回収が求められる。Jambeckらによると実際に、毎年880万トンのプラスチックが海に放流されている [8]。プラスチックを含む海岸の漂着ゴミは早急な回収が求められている一方で、回収活動に伴い発生する問題もある。

2.1 海岸のゴミ問題

ゴミが多く漂着する離島での回収活動では、回収人員の確保や、人を乗せるための船の用意が必要となる。船に関しては、回収に必要な人員や道具などに加えて、回収したゴミも乗せる必要がある。したがって、回収活動を行う際には、事前に海岸のゴミの量を想定して適切な数の人員や船を用意しなければならない。

海岸の漂着ゴミのデータとしては、環境省が公表しているデータが存在する [14]。全国の代表的な海岸で漂着ゴミの調査を行なった際の手法や結果などを示すものである。しかし、これらは年度単位など一定期間でのデータになる。そのため、実際に回収活動を行う際に、例えば、台風があった翌日など、ゴミの量に短時間で大きな変化があった場合には適切に把握することができない。また、全国の海岸を網羅している訳ではないため、全国各地で行われる回収活動に有効活用できない。環境省のデータだけでは回

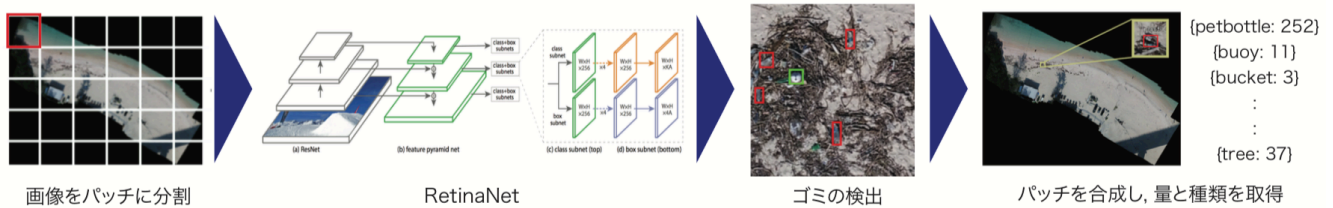


図 1 Ortho-detector の構成図：分割した画像を RetinaNet の入力とし、ゴミの種類と量に関する情報を定量的に把握する。

取活動の事前準備の段階で、海岸全体のゴミの種類と量を適切に把握するには限界がある。海岸ごとのゴミの種類と量を正確に把握できるようなデータが存在しないのが現状である。

一方で、ゴミの総量を把握できないまま、人員や船の用意をすることは容易ではない。船の用意などの回収活動に伴うコストが高いことから、そもそも回収活動自体が頻繁に行える状況ではないため、回収すべきゴミの量を上回る数の人員や船を用意するようなことは避ける必要がある。そのため、清掃計画の策定の段階で、実際に回収活動を行う時や場所に関する適切な情報を得られるデータが求められる。

2.2 関連研究

地上から漂着ゴミをセンシングする方法はいくつか存在する。Mirco らは Frame 法と Rake 法と呼ばれる二種類のサンプリング法で地上から海岸に落ちているゴミを採集し、分析した [7]。Frame 法では、海岸をフレームごとに分割し、ゴミを採集し、ふるいにかけて分類したのに対して、Rake 法では、海岸線から崖などのビーチの幅全体をカバーし、海岸に近い部分、草の生えている部分、などの領域毎で分割して、採集・分析することで、海岸のゴミの種類と量を示している。

また、Adbullah らは beachbot というロボットを VR で人間が操作することにより、地上からのゴミ回収を試みている [1]。これは、センシングを念頭においたものではなく、VR 技術を活用してロボットを操作することで楽しくゴミの回収ができるというものであるが、海岸を動き回るロボットを使用すれば、地上からのアプローチであってもセンシングを自動化することができる。しかし、通常海岸は広いため、漂着ゴミの正確な位置や量の把握に時間がかかることに加え、地上から広範囲にわたる海岸のゴミの種類と量を把握することは、いずれの方法でも難しい。beachbot のようなロボットを用いても、岩場や水たまりの多い海岸では、地上から正確なセンシングを行うことは難しいため、漂着ゴミのセンシングには上空からのセンシングが必要となる。

近年は、ドローンが手軽に入手できるようになったこともあり、以前よりも上空からの広範囲なセンシングが容易

にできるようになった [5]。Zhongcong らは、ドローンを用いて撮影した海岸の画像を RGB のそれぞれの波長の反射率の違いに着目し閾値を設定して分析することで、ゴミの密度を推定した [2]。しかし、この手法ではゴミを種類ごとに特定することは難しく、ゴミの場所や種類を適切に測ることはできていない。また、海岸ごとのゴミの特性などに合わせて色の閾値を変える必要もある。

さらに、深層学習と画像処理技術の発展に伴い、画像から特定の物体の検出を高い精度でできるようになった。ドローンで撮影した画像に対して物体検出を行うことで、空中から地上の物体を検出することも可能になり、建物や船などの大きいものの検出だけでなく、車などの検出が成功している例もある [3], [6], [11]。

3. Ortho-detector

本研究では、ドローンで撮影された海岸のオルソ画像から、物体検出技術を利用してゴミを種類毎に検出し、種類と量を定量的に把握する。Ortho-detector では検出に使用する画像を、1) 分割、2) 各パッチに対して検出、3) 結合することでパッチ毎に処理をする。図 1 に全体の流れを示す。

3.1 オルソ画像の分割

検出に使用する画像は、海岸の広範囲を撮影したもので、一枚のオルソ画像は平均すると約 12000 × 21600 ピクセルにもなるため、GPU を用いても一度に処理をすることは難しい。そこで、まず、入力された画像を 600 × 600 ピクセルのパッチに分割する。モデルの予測結果が漏れ無く画像に反映されるよう、被りを持たせて分割する必要があるため、480 ピクセルずつずらして分割を行う。

3.2 漂着ゴミの検出

次に分割した各パッチに対して検出を行う。海岸全体を写した画像では、検出対象となるゴミの部分よりも砂浜や海、岩場などの背景に当たる部分が圧倒的に多いため、検出対象とする物体と背景部分とを明確に区別する必要がある。そこで、背景と検出物体との区別ができ、衛星画像に対しても高い精度の検出が可能な RetinaNet というモデルを採用した [9]。今回使用するモデルは、物体検出に利用す

る一般的なデータセットである MSCOCO データセットで事前学習させたモデルに対して、追加学習を行ったものである [10].

3.3 海岸ゴミマップ作成

最後に、パッチ間で重複しているエリアを一つにまとめる。各パッチの検出結果を矩形としてオルソ画像に描画することで、海岸ゴミマップを作成する。被りを持たせて分割した画像を結合するため、結合するパッチそれぞれで予測結果が出力され、矩形が多数表示される可能性がある。そのため、重なっている矩形の組に対して、それぞれに付与されている確率が高い方を採用する、Non Maximum Suppression (NMS) をモデルの出力後に加えて、パッチをまとめた後にも追加した。また、海岸のゴミは密集していることもあり、モデルが学習する時に重なりのある矩形を考慮できるように、モデルの出力後とパッチ結合後のいずれの NMS に関しても、NMS のオーバーラップの割合は 0.15 とし、予測後も重なりのある矩形をある程度許容するように設定した。

4. オルソ画像データセット

漂着ゴミを検出するにあたり、有効活用できるデータセットがなかったため、独自のデータセットを作成した。本章では、オルソ画像データセットに使用した画像の撮影方法とオルソ画像の作成方法、検出対象のアノテーションについて述べる。

4.1 海岸撮影

海岸の撮影はドローンを使用して行う。撮影場所は、沖縄県の久手堅海岸 (Kudeken)、真喜屋運動広場裏の海岸 (Makiya)、米須海岸 (Komesu) である。様々な天候条件が考慮されるが、ドローンが飛行できない雨天時を除いた日に撮影を行なった。撮影高度は第 5 章の結果から 20m で統一する。この撮影高度はドローンを飛ばした地点からの距離になるため、同じ海岸の画像であっても場所によって厳密な高度は変わる可能性がある。使用するドローンは Inspire2 で、カメラは DJI Zenmuse X5s、レンズは MFT15mm f/1.7 ASPH である。20m の撮影高度でレンズが 15mm のものを使用すると、撮影した画像の 1 ピクセルに対応する地上での距離を表した値である、Ground Sample Distance (GSD) は 0.4cm/px となる。今回は、撮影する際にドローンのコントローラの画面に表示される GSD の値をそのまま使用した。

海岸全体を撮影するためのドローンの飛行ルートは、DJI GS PRO という iOS アプリケーションで指定した。撮影中は、ドローンが自動で移動しながらその範囲を動き、部分ごとにホバリング撮影を行う。各地点でドローンによって撮影される画像のサイズは 5280 × 3956 ピクセルである。

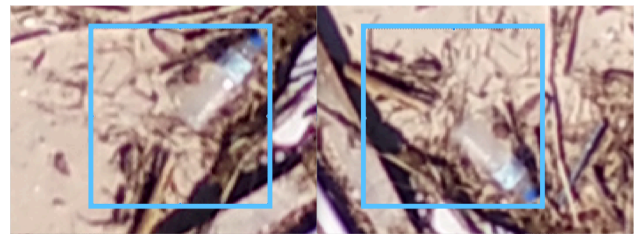


図 2 Pix4d で作成したオルソ画像の解像度の比較：15m で撮影された元画像 (左) とオルソ画像 (右) による解像度の比較結果。青枠は 100 × 100 ピクセルを表している。

この時、撮影範囲全体を網羅するべく、ドローンが移動する際には、隣接する地点間の画像で重複するように撮影する。隣接する地点間の重複度合いは縦 80%、横 50% の割合で設定した。

4.2 オルソ画像作成

オルソ画像を使用し、物体の写り方を統一することで、物体を検出するタスクを簡単にできる。また、海岸全体を写したオルソ画像には、同じ物体が複数回写ることがないため、漂着ゴミの量の測定が容易になる。

ホバリングしながら撮影した画像は、海岸の部分ごとの大量な画像になるため、それらをまとめて海岸全体のオルソ画像に変換する。空中画像からオルソ画像に変換する際には、Pix4d を使用した。Pix4d を用いてオルソ化することにより、画像の補正が行われることによって多少のノイズが生じる場合があるものの、実際にドローンで撮影した元データと、オルソ画像に変換したものを比較するとオルソ画像では、解像度が約 1.1 倍になった。図 2 では、15m からドローンで撮影した元画像 (左) とオルソ画像 (右) で、同じ部分を表示したものである。100 × 100 ピクセルを青い枠で囲ってある。

4.3 検出対象

データセットでは、ペットボトル・ブイ・バケツ・網・縄・流木・スプレー缶をアノテーションする。有害性のあるプラスチック類のほか、オルソ画像から確認できるものをアノテーション対象とした。アノテーションは VoTT を用いて対象物体を矩形で囲うようにして行った [12]。物体全体がはっきりと写っていないものにはアノテーションをしなかった。表 1 では、Kudeken, Makiya, Komesu の 3 つの海岸ごとのアノテーション数を表している。それぞれの海岸のオルソ画像の画像サイズは、Kudeken が約 18000 × 40200 ピクセル、Makiya が約 7800 × 11400 ピクセル、Komesu が約 10200 × 12000 ピクセルである。表中の pb は plastic bottle を、can は spray can をそれぞれ表している。ブイなどの漁具は実在する数が少なかったが、スプレー缶などの小さなものは 20m で撮影した画像では、物体を特定することが困難であり、アノテーション数は極

coast	tree	rope	pb	net	can	bucket	buoy	sum
Kudeken	10	3	26	3	0	0	0	42
Makiya	64	13	168	4	0	1	2	252
Komesu	6	3	23	0	0	0	1	33
sum	80	19	217	7	0	1	3	327

表 1 作成したデータセットの各海岸に置けるアノテーションされた物体数：pb は plasticbottle を，can は spraycan を表している。

端に少なくなっている。全ての海岸に共通して存在していたのはペットボトルと流木と縄だけであった。

5. 撮影高度の検証

ゴミの種類や位置，量の推定を行う前に，事前実験として，撮影高度と検出精度を比較して検出に適切な撮影高度を調べる。ドローンで撮影した画像は高度が低くなればなるほど解像度が上がり，Ortho-detector は学習しやすくなる。一方で，高度を下げると Ortho-detector の学習精度はよくなるものの，ドローンの移動距離は長くなり，撮影に時間がかかるだけでなくバッテリーの消費量も多くなる。そこで，学習精度を保てて，かつ撮影高度がより高いものを最適な撮影とし，より効率良く撮影・実験を行うために，最適な高度の比較実験を行う。本章では，撮影高度の比較実験の際の実験設定とその結果について述べる。

5.1 実験設定

ドローンの撮影高度を検証するためには，Ortho-detector が学習できていることをしっかりと確認する必要がある。そこで，高度毎で小規模なデータセットを作成し，学習とテストで同一の画像を用いることで Ortho-detector が学習できているかを確認する。検証する高度は，5m から 30m までの 5m 間隔とする。画像は 4 章で作成したオルソ画像のデータセットに含まれる Komesu を使用する。高度毎に撮影したものを使用し，それぞれの高度で作成したオルソ画像を 600×600 ピクセルに分割した画像から 12 枚を選んでデータセットとする。検出対象は，小さくて色が透明で背景と同化してしまいやすく，検出が難しいペットボトルのみとし，データセットを作成した。なお，高度が低ければ低いほど，使用する画像中にペットボトルがあまり写っていないということが起こりうるため，データセットに含まれる 12 枚の画像は隣接しているものではなく，ペットボトルがより多く写っているものを選ぶようにした。

次に，作成した小規模データセットを使用して，Ortho-detector の物体検出の部分だけを利用して実験を行なった。Komesu の小規模データセットを入力画像として，Ortho-detector に入力した後，分割した画像を合成し，一枚の画像に見立てて可視化する。今回の事前実験では，隣接している画像を使用していないため，一つにまとめる必要はないが，実際に実験を行う際は，海岸の全体の画像に対して Ortho-detector の予測結果を表示するため，事前実験でも



図 3 20m での予測結果：Komesu のうちペットボトルがある 12 枚の画像でテストを行なった可視化画像を 600×600 で切り取った。赤枠で囲まれているものがペットボトルである。

GSD(cm/px)	0.1	0.2	0.3	0.4	0.5	0.6
Altitude(m)	5	10	15	20	25	30
Plastic bottles	17	16	19	19	18	22
Precision	1.00	1.00	1.00	1.00	1.00	1.00
Recall	1.00	1.00	1.00	1.00	0.84	0.30
AP	1.00	1.00	1.00	1.00	0.84	0.30

表 2 撮影高度と検出精度の比較結果：Plastic bottles はアノテーションの数を表している。20m 以下で撮影した画像では，Precision, Recall, AP の値が全て 1 になった。

同様の方針で行なった。

評価指標としては，物体検出の一般的な指標である，Precision, Recall, Average Precision (AP) を採用する。学習とテストを同一の画像で行うことから，これら全ての値が 1 にならないと Ortho-detector がしっかりと学習できているとは言えないことになり，実験に有効な高度ではないということになる。

5.2 結果

表 2 では，各行で GSD, GSD に対応するドローンの撮影高度，画像中のペットボトルの数，そして 3 つの評価指標を，列では GSD (撮影高度) ごとのペットボトルの数と評価指標を定量的に表している。Precision, Recall, AP の全てのスコアが 1 になったのは 20m 以下であった。また，定性的結果として，図 3 では 20m での Ortho-detector の予測結果を可視化した。600×600 ピクセルに分割したものを一枚の海岸の画像に見立てて出力している。

5.3 Makiya を使用して検出

撮影高度の検証実験では，モデルの学習がしっかりと行われていることを確認するため，小規模なデータセットを

Annotations	168
Images	53
Precision	0.88
Recall	0.97
AP	0.91

表 3 Makiya を使用した場合の検出結果: Annotations はペットボトルのアノテーション数を, Images が 600 × 600 ピクセルの画像数を表している. 90%以上のペットボトルを検出できた.

用いたが, より実践的なタスクを想定して, データセットの数を増やし, 20m で撮影した画像を使用して検出を行う. 使用したデータセットは Komesu よりもアノテーション数の多い, Makiya からペットボトルのアノテーションがある 53 枚を使用した. 撮影高度の検証実験と同様に, 学習に使用した画像と同じ画像を用いてテストを行う. 表 3 は学習結果を示しているが, データセット内に含まれるペットボトルは 168 個であったが, そのうち, 90%以上のゴミを検出することができた.

6. 実験

事前実験の結果を踏まえて, 4 章で作成したデータセットを使用して Ortho-detector の学習, 及び漂着ゴミの検出を行う. 本章では, 実際の用途として, Ortho-detector の学習に使用した海岸と同じ海岸でゴミの検出を行う場合を想定して実験を行い, その結果を述べる. 次に, 検出を適切に評価する際に行なったパラメータの調整や, また Ortho-detector の学習に使用した海岸とは異なる海岸で予測を行う場合を想定した実験とその結果に付いて述べた後, 全体の考察を行う.

6.1 実験設定

本実験では, Ortho-detector の学習に使用した海岸と同じ海岸で予測を行う場合を想定して, 実験を行う. 使用する画像は, オルソ画像データセットからアノテーション数が一番多い Makiya を用いる. 同じ海岸での検出を想定するため, Makiya の画像を学習用とテスト用に 9:1 でランダムに分け, 学習用データセットだけを用いて Ortho-detector を学習し, 学習に使用していないテストデータセットで予測を行い AP を用いて評価する. 検出対象は, オルソ画像データセット内のどの海岸でもアノテーションされていて, その他の海岸でも確認できると予想される, ペットボトル・流木・縄の 3 つとする. 検出対象の 3 つの物体のいずれかのアノテーションがある画像のみを使用して, 学習用データセットとテスト用データセットを作成した. 3 つのクラスのうち, いずれかのアノテーションがある画像は, Ortho-detector を用いて分割した 600 × 600 ピクセルのパッチが 70 枚であったため, 63 枚を学習に, 7 枚をテストに使用した. しかし, データセット内の全ての画像に



図 4 3 つのクラスで学習したモデルの検出結果: 緑で囲まれているものがペットボトルで, 赤で囲まれているものが流木.

class	plasticbottle	tree	rope
AP	0.46	0.13	0.10

表 4 Makiya を 9:1 で学習用とテスト用のデータセットに分割して予測を行なった結果.

対して同じ数のアノテーションがされているわけではないため, 画像の分け方によって学習精度が変化することが懸念される. そのため, データセットをランダムに分割して学習・テストをすると行った行程を 5 回繰り返す, それぞれで得られた AP の平均をとってテストの結果とした.

また, 検出クラスの数に検出精度に影響を与えるか比較するため, アノテーション数が一番多いペットボトルだけでも同様の実験を行なった. ペットボトルのみの場合, アノテーションのある画像は全部で 53 枚であったため, 47 枚を学習に, 6 枚をテストに使用した.

6.2 結果

まず, 3 つのクラスを検出した場合の平均の AP は, ペットボトルで 0.46, 流木で 0.13, 縄で 0.10 となった (表 4). 図 4 は, 定性的な結果として 3 つのクラスで検出を行なった画像を可視化したものの一部分である. 表 1 に示した通り, クラス毎のアノテーション数はペットボトルで 168, 流木で 64, 縄で 13 と大きな差があるため, アノテーション数の少ない流木や縄はうまく学習できなかった.

また, ペットボトルのみを対象に学習と予測を行なった場合の平均の AP も 0.46 となり, 検出するクラス数が減ったことによる影響はなかった.

6.3 評価指標のパラメータ設定

次に, 海岸の漂着ゴミ検出タスクを分析するために, 検出対象をペットボトルのみに絞って, 追加実験を行った.

第5章や本章で述べている通り、物体検出の評価指標としては、Precision, Recall とそこから計算される AP を使用している、これらの評価指標は、学習したモデルから予測される矩形がどのくらいアノテーションデータと合っているかを計算したものである。モデルが予測した矩形は膨大な数になるため、評価に使用する矩形を限定するような制約が必要になる。そこで、NMS のオーバーラップの割合、画像一枚あたりの矩形の最大許容数といったパラメータの調整を行なった。モデルが学習できていることを確認するため、Makiya の全ての画像を使用して学習を行い、同じ画像で予測を行なった。評価を行う際の NMS のオーバーラップの割合は 0.9 とし、同じ物体に対して何個も矩形が出現しないよう制約するため大きな値にした。

また、画像ごとの矩形の数そのものを制約する必要があるが、Makiya では全ての画像に同じ数のアノテーションがあるというわけではないため、アノテーションのない画像では誤検出が多く発生しうる。例えば Makiya の場合、アノテーションが1個もない画像から最大で13個付いている画像もある。画像1枚あたりの矩形の最大許容数を小さく設定してしまうと、検出漏れが生じてしまう。一方で、大きな値に設定すると、アノテーションが少ない画像に対する誤検出の件数が多くなるため、最適なパラメータの設定が重要になる。今回 Makiya では、画像1枚あたりの最大許容数をアノテーション数の最大値である13に設定したところ、Precision は 0.82, Recall は 0.97, AP は 0.83 となった。

6.4 学習に使用していない海岸で予測

追加実験として、学習に使用した画像と違う海岸の画像で予測を行う場合を想定した実験を行なった。これまで、Makiya で学習した Ortho-detector を使用して Makiya でテストをしていたが、実際には離島に頻繁に行くことは困難であることから、撮影に行きやすい本島の海岸で撮影した画像で学習をして、離島の海岸でテストをするといったことが考えられうる。そのため、Makiya で学習した Ortho-detector を Kudeken と Komesu で予測することで他の海岸への予測結果とする。学習に使用したモデルは、第5章の Makiya を使用して検出を行なったものをそのまま使用して、ペットボトルのみを対象にして検出を行なった。テストに使用した画像は Kudeken と Komesu のうち、ペットボトルのアノテーションがある部分だけに限定したところ、Kudeken と Komesu それぞれで、600×600 の画像は 23 枚と 15 枚、ペットボトルのアノテーションは 26 個と 23 個になった。Kudeken では AP が 0.50 だったのに対して、Komesu では 0.19 と低い結果になった(表5)。Kudeken と Komesu でいずれも Recall はあまり変わらず、Precision に差があったことから Komesu の予測結果が悪かった原因として、誤検出が多かったことがわかる。

coast	Kudeken	Komesu
Annotations	26	23
Images	23	15
Precision	0.68	0.26
Recall	0.57	0.54
AP	0.50	0.19

表5 Makiya で学習したモデルで異なった海岸の漂着ゴミを予測した結果: Annotations はデータセットに含まれるアノテーションの数を、Images は 600×600 ピクセルの画像数を表している。

6.5 全体の考察

Kapli らが衛星画像に対して物体検出を行なった結果の AP は 0.77 であった [13]。Makiya での検出では AP は 0.46 であったが、これは物体検出のタスクとしては低すぎる訳ではない。しかし、Kapli らの結果と比較すると、データセットが違うものの、Ortho-detector にはまだまだ改良の余地はあると考える。

検出精度を上げていくためには、十分なデータセットが必要になる。Arthur らも同様に、空中画像から物体検出を行なっているが、使用したデータセットには 1000 件を超えるアノテーションが含まれている [4]。そのため、今後は物体ごとに最低でも 1000 件ほどのデータが必要になるだろう。

また、クラスそのものに関しても、データが増えることで検出に使用するものが変わってくる可能性がある。今回はペットボトル、流木、縄のみを検出に使用したが、データセットが増えることで、実際には他の物体が多く発見されるようであれば、検出対象は変える必要がある。それに加えて、より多くの海岸を撮影し、データセットを作成していく段階で、漂着ゴミのうち、多く存在している種類が明白になることから、学習や検出に使用するクラスの定義そのものも変わる可能性がある。

追加実験で異なる海岸でテストした際に、検出結果にばらつきがあった原因としては海岸の砂の質の違いや天候が考えられる。Makiya が黒っぽい砂であるのに対して、Kudeken は白浜、Komesu では少し黄みがかかった白浜であった(図5)。また、Makiya や Kudeken のペットボトルでは、光の反射があまり見られなかったのに対して、Komesu のペットボトルの端には、太陽光の反射が見られる。天候が異なることで、物体の写り方も異なってくるため、光の具合を調整するなどの工夫が今後必要になる。データセットの確保や天候を考慮して光の具合を調整することで、異なる海岸に対しての汎化性能も上げていくことも今後の課題としたい。

また、追加実験で行なった評価指標に関するパラメータの設定は手動で行なったが、現時点で最適であるかどうかは断言できない。パラメータの調節は手動でやるには限界があるので、パラメータの値を自動で調整するといったこ



図 5 オルソ画像データセット内の画像。左から, Makiya, Kudaken, Komesu. それぞれの画像でペットボトルのある部分を 200 × 200 ピクセルで切り取った。

とも検討する余地があるだろう。

7. まとめ

本研究では, 海岸のオルソ画像から物体検出技術を利用して, ゴミを種類ごとに検出することで種類と量を定量的に把握する手法として Ortho-detector を提案した。また沖縄の海岸のオルソ画像を使用し, 実際の用途を想定して検出タスクの分析を行なった。結果としては, モデルの学習に使用した海岸と同じ海岸での予測では, AP が 0.46 となった。異なる海岸で予測を行なった場合には検出がうまくできなかったものもあり, データセットの確保はもちろんのこと, 海岸の砂の色や天候などをいかに考慮するかが今度の課題である。

謝辞 データを提供して頂いた株式会社紘の森に感謝を申し上げます。

参考文献

- [1] Almosalami, A., Jones, A., Tipparach, S., Leier, K. and Peterson, R.: Beachbot: crowdsourcing garbage collection with amphibious robot network, *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ACM, pp. 333–334 (2018).
- [2] Bao, Z., Sha, J., Li, X., Hanchiso, T. and Shifaw, E.: Monitoring of beach litter by automatic interpretation of unmanned aerial vehicle images using the segmentation threshold method, *Marine pollution bulletin*, Vol. 137, pp. 388–398 (2018).
- [3] Boonpook, W., Tan, Y., Ye, Y., Torteeka, P., Torsri, K. and Dong, S.: A Deep Learning Approach on Building Detection from Unmanned Aerial Vehicle-Based Images in Riverbank Monitoring, *Sensors*, Vol. 18, No. 11, p. 3921 (2018).
- [4] Douillard, A.: Object Detection with Deep Learning on Aerial Imagery, <https://medium.com/data-from-the-trenches/object-detection-with-deep-learning-on-aerial-imagery-2465078db8a9>.
- [5] Eling, C., Klingbeil, L. and Kuhlmann, H.: Real-time single-frequency GPS/MEMS-IMU attitude determination of lightweight UAVs, *Sensors*, Vol. 15, No. 10, pp. 26212–26235 (2015).
- [6] Gallego, A.-J., Pertusa, A. and Gil, P.: Automatic ship

classification from optical aerial images with convolutional neural networks, *Remote Sensing*, Vol. 10, No. 4, p. 511 (2018).

- [7] Haseler, M., Schernewski, G., Balciunas, A. and Sabaliuskaite, V.: Monitoring methods for large micro-and meso-litter and applications at Baltic beaches, *Journal of coastal conservation*, Vol. 22, No. 1, pp. 27–50 (2018).
- [8] Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., Narayan, R. and Law, K. L.: Plastic waste inputs from land into the ocean, *Science*, Vol. 347, No. 6223, pp. 768–771 (2015).
- [9] Lin, T.-Y., Goyal, P., Girshick, R., He, K. and Dollár, P.: Focal loss for dense object detection, *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988 (2017).
- [10] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L.: Microsoft coco: Common objects in context, *European conference on computer vision*, Springer, pp. 740–755 (2014).
- [11] Razakarivony, S. and Jurie, F.: Vehicle detection in aerial imagery: A small target detection benchmark, *Journal of Visual Communication and Image Representation*, Vol. 34, pp. 187–203 (2016).
- [12] Team, M. P. C.: Visual object tagging tool, <https://github.com/microsoft/VoTT>.
- [13] Varshtey, K.: Object Detection On Aerial Imagery Using RetinaNet, <https://towardsdatascience.com/object-detection-on-aerial-imagery-using-retinanet-626130ba2203>.
- [14] 環境省: 平成 29 年度漂着ごみ対策総合検討業務報告書, <http://www.env.go.jp/water/marine.litter/hyouchaku-gaiyouban.pdf>.