

S3-3: 河川のDX, ビッグデータ分析, AI

ドローンによる河川巡視の自動化を見据えた、 河川敷に点在する投棄物の自動検出手法に関する検討

○ (株)東京建設コンサルタント 小島崇
岡山大学大学院 横田翔一
岡山大学准教授 吉田圭介



河川監理の高度化・河川投機物の検出技術

ドローン+AI

藤井純一郎, 吉田龍人, 天方匠純, 山下隆義

「UAV写真測量と深層学習を活用した河川管理手法の研究」

AI・データサイエンス論文集 1 (J1), 588-595, 2020-11-11, 2020.

Semantic Segmentationを用いた河川地被分類

- ・ 撮影高度の違い・オルソ化による分類精度の評価

岡野将大, 吉田龍人, 高橋悠太, 藤井純一郎, 天方匠純

「河川の不法投棄物検知におけるマルチモーダルモデルの適用検証」

The 37th Annual Conference of the Japanese Society for Artificial Intelligence, 2023.

マルチモーダルモデルを用いた河川ゴミの検出

- ・ グリッド上に分割した画像内にゴミの有無＋分類
- ・ AIモデルの違いによる精度検証

門脇大典, 中谷祐介, 甲斐達也

「生成AIによる学習データ拡張と河川浮遊ゴミ検出モデルへの応用」

土木学会論文集, Vol. 81, No. 10, 25-00172, 2025.

生成AIを用いた教師データの作成と物体検出を用いた河川浮遊ゴミの検出

- ・ Stable Diffusionによる河川浮遊ゴミ画像の生成
- ・ Yoloを用いた河川浮遊ゴミ検出

河川監理の高度化・河川投機物の検出技術

著者らの先行研究

吉田圭介, 谷口純一, 潘是均, 小島崇, 永田貴美久, 渡邊明英
「ALB点群データを援用した河川航空写真の深層学習による地被分類法の検討」
土木学会論文集B1(水工学) Vol.76, No.2, I_1003-I_1008, 2020.

Semantic Segmentationを用いた河川地被分類

- ・ ALBから得られた樹木密度, 樹高データとRGB可視画像を組み合わせたモデル
- ・ 竹林と草本, 大本の分類

下江大地, 潘是均, 吉田圭介, 西山哲, 小島崇
「UAVを活用した河川巡視時のペットボトル検出モデルにおける画像生成AIの応用」
土木学会論文集, Vol. 81, No. 16, 24-16180, 2024.

生成AIを用いた教師データの作成と物体検出を用いた河川ゴミ(ペットボトル)の検出

- ・ Stable Diffusionによる河川に投棄されたゴミ(ペットボトル)画像の生成
- ・ Yoloを用いた河川ゴミ検出

英佑成, 吉田圭介, 潘是均, 小島崇, 西山哲
「UAV写真と大規模マルチモーダルモデルを用いた河川投棄物検出手法の高度化に関する研究」
土木学会論文集, Vol. 82, No. 16, 25-16082, 2026.

物体検出モデルとマルチモーダルモデルを組み合わせたゴミ検出手法の提案

- ・ Yoloによるゴミの検出結果をLLaVAを用いて種類を分類
- ・ 撮影高度による推論精度の比較

河川投機物の検出に関する課題

1) ゴミの種類や状態はさまざま

- ・ 缶やペットボトル, ゴミ袋, 段ボール, 発泡スチロールなどを想定しているが, 粗大ごみや紙類, 樹木等様々なゴミが存在している.
- ・ 投棄されてから河川を浮遊・流下するものや、土砂や草本に埋もれるなど, 完全な状態でないものも多い.

教師ありモデル（物体検出モデル）によるゴミの検出には限界がある.
事前に様々なゴミのデータを学習させることが難しい.

2) 撮影高度による精度への影響

- ・ 効率的かつ安全にUAVの巡視を行うためには, 十分な高度が必要.
- ・ 鮮明に撮影するためには,  低高度で撮影する事が望ましい.

高度によってゴミの見え方や解像度にも大きく差が出る.

3) 解析環境

- ・ 巡視によって得られた高解像度の画像をローカル環境(PC)で処理したい.

推論コストとデータの保管上の問題

UAVオルソ画像への適用

- ・ 適用対象
- ・ 検出するゴミ



適用対象（草地,階段）

岡山県旭川

高度10m

13,376 × 10,688pixel

高度30m

3,932 × 3,322pixel

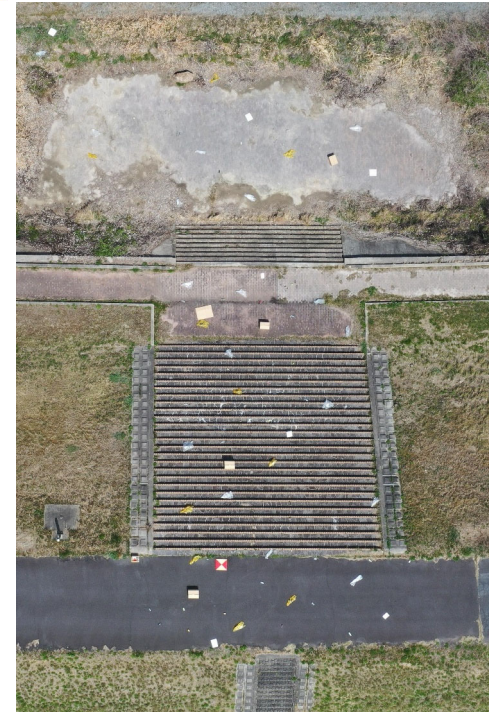
高度60m

2,064 × 1,593pixel

草地



階段



9,424 × 12,464pixel

2,992 × 4,688pixel

1,894 × 2,807pixel

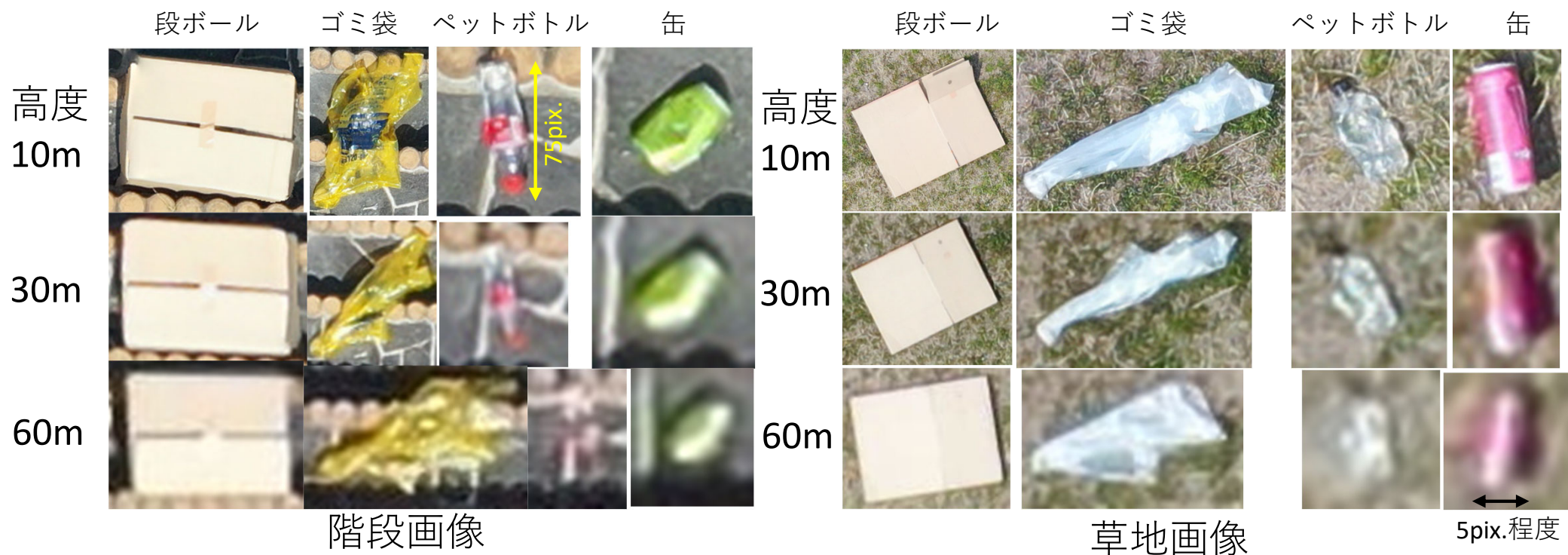


株式会社 東京建設コンサルタント
TOKEN C. E. E. Consultants Co., Ltd.

適用対象（高度別のゴミの見え方）

河川敷に配置したゴミの種類

ゴミの種類	缶	段ボール	ゴミ袋	ペットボトル	発砲スチロール	その他
プロンプト上の表現	Can	Cardboard	Plastic bag	Plastic bottle	Styrofoam	Othrrers
長辺の寸法	10～15cm	60cm	70～90cm	15～20cm	25～40cm	40～60cm



缶やペットボトルは高度60mでは、短辺が数pix. 程度になる。
 高度60mではオルソ化によって画像が不鮮明になり、輪郭等がぼやける。

AIモデル単体での評価

- Object Detection
- VLM
- Autoencoder(AE)



AIを用いた検出(1)

Object Detection

Ultralytics社のYoloを用いた。
教師データ：AIGC

先行研究では、Yolo v5を用いていた。
本研究では、新しいv8や11, 26を適用し
モデルやサイズによって精度がどの様に
変化するかを確認し、
物体検出モデルの課題を確認した。

先行研究から引用

下江大地, 潘是均, 吉田圭介, 西山哲, 小島崇 :
UAV を活用した河川巡視時のペットボトル検出モデルにおける画像生成 AI の応用,
土木学会論文集, Vol. 81, No. 16, 24-16180, 2024.

表-1 AIGC (Stable Diffusion, txt2img)での
画像生成時の条件設定

項目	入力値または内容
Prompt	Bikes, Cardboards, PET, Plastic Bags, High resolution, 8k camera, Bird view, UAV view, Riparian area
Sampling step	150
Width/Height(pix)	608/608, 900/600
Batch Size/Count	1/100
CFG Scale	7



図-4 AIGCの画像の一例

モデル/サイズと検出精度の比較

		草地									階段								
		10m			30m			60m			10m			30m			60m		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Yolov8	n	0.210	0.731	0.326	0.750	0.570	0.648	0.795	0.403	0.535	0.313	0.733	0.439	0.814	0.402	0.539	0.870	0.241	0.377
	x	0.299	0.667	0.413	0.831	0.747	0.787	0.783	0.468	0.585	0.243	0.570	0.340	0.870	0.460	0.602	0.833	0.241	0.374
Yolo11	n	0.290	0.833	0.431	0.798	0.848	0.822	0.810	0.610	0.696	0.339	0.756	0.468	0.786	0.506	0.615	0.923	0.289	0.440
	x	0.340	0.680	0.453	0.806	0.684	0.740	0.793	0.546	0.646	0.355	0.709	0.473	0.800	0.506	0.620	0.946	0.422	0.583
Yolo26	n	0.311	0.590	0.407	0.750	0.342	0.470	0.667	0.026	0.050	0.366	0.733	0.488	0.968	0.345	0.509	0.966	0.337	0.500
	x	0.278	0.692	0.397	0.831	0.684	0.750	0.800	0.416	0.547	0.127	0.256	0.170	0.882	0.172	0.289	0.882	0.181	0.300

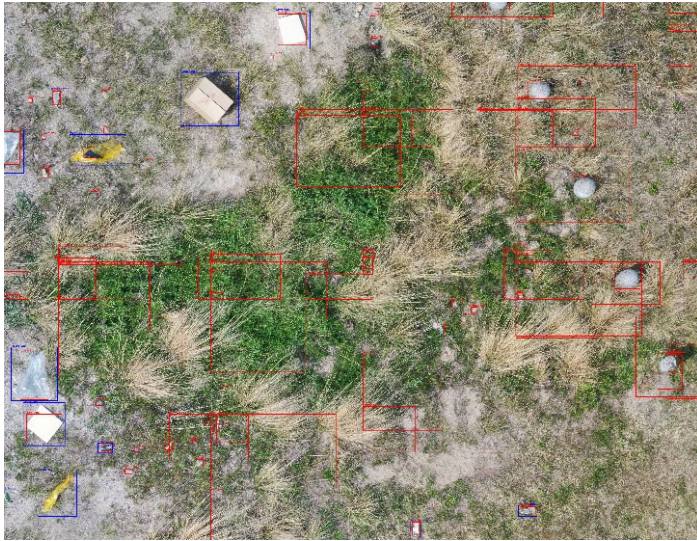
- ・ Yolo11n:30mが高い精度で検出しているのに対して、Yolo26xはPrecisionが低く検出精度が低い。
- ・ 草地と階段では、階段は全体的に検出精度が低い傾向にあった。

Conf.=学習時のF1が最適となる値を採用



AIを用いた検出(1)

草地
10m



Object Detection

— 正解
— 物体検出

- ・ 草地10mでは、草をゴミと誤検出している。
- ・ 缶の検出はできているが、黄色いゴミ袋の検出ができていない。

草地
30m



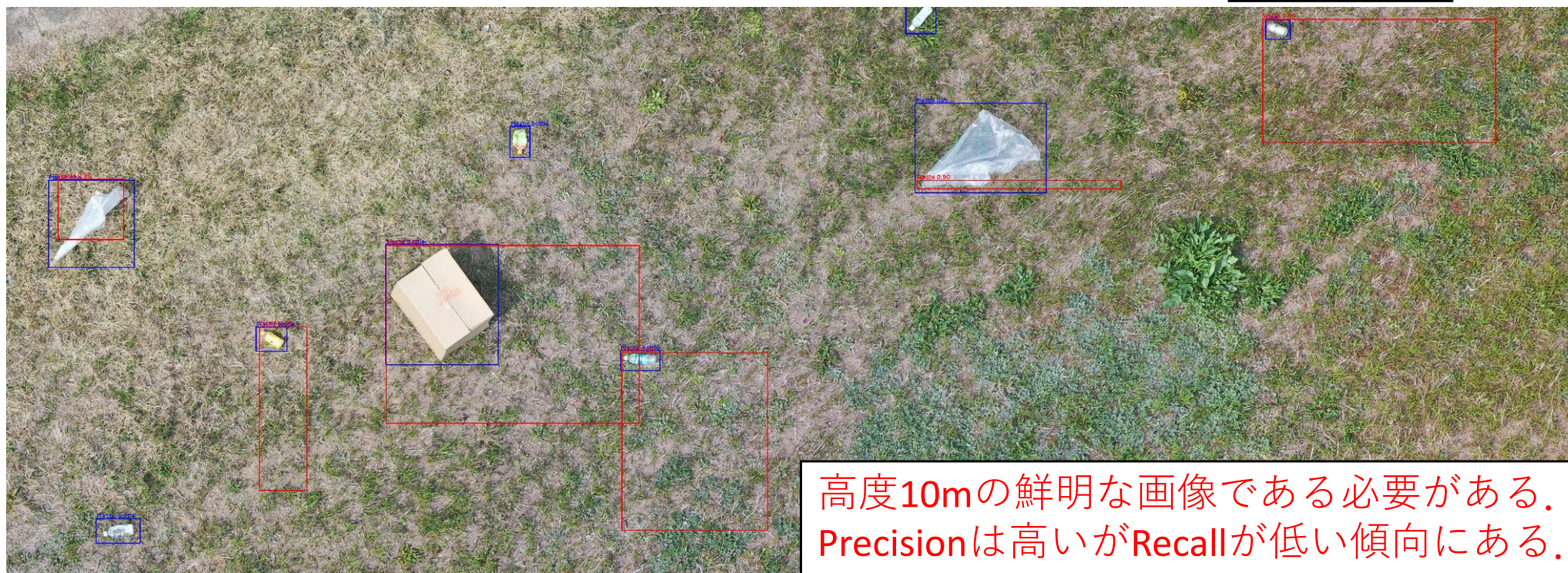
- ・ 草地30mでは、缶やペットボトルの検出ができていない。
- ・ 段ボールやゴミ袋は概ね検出されている。

教師ありモデルでは、学習したゴミの解像度の影響から、解像度が低下すると缶やペットボトルの様な小さいゴミは検出できなくなる

AIを用いた検出(2)

VLM

— 正解
— VLM



高度10mの鮮明な画像である必要がある。
Precisionは高いがRecallが低い傾向にある。

草地:10m, Qwen3-vl:8b

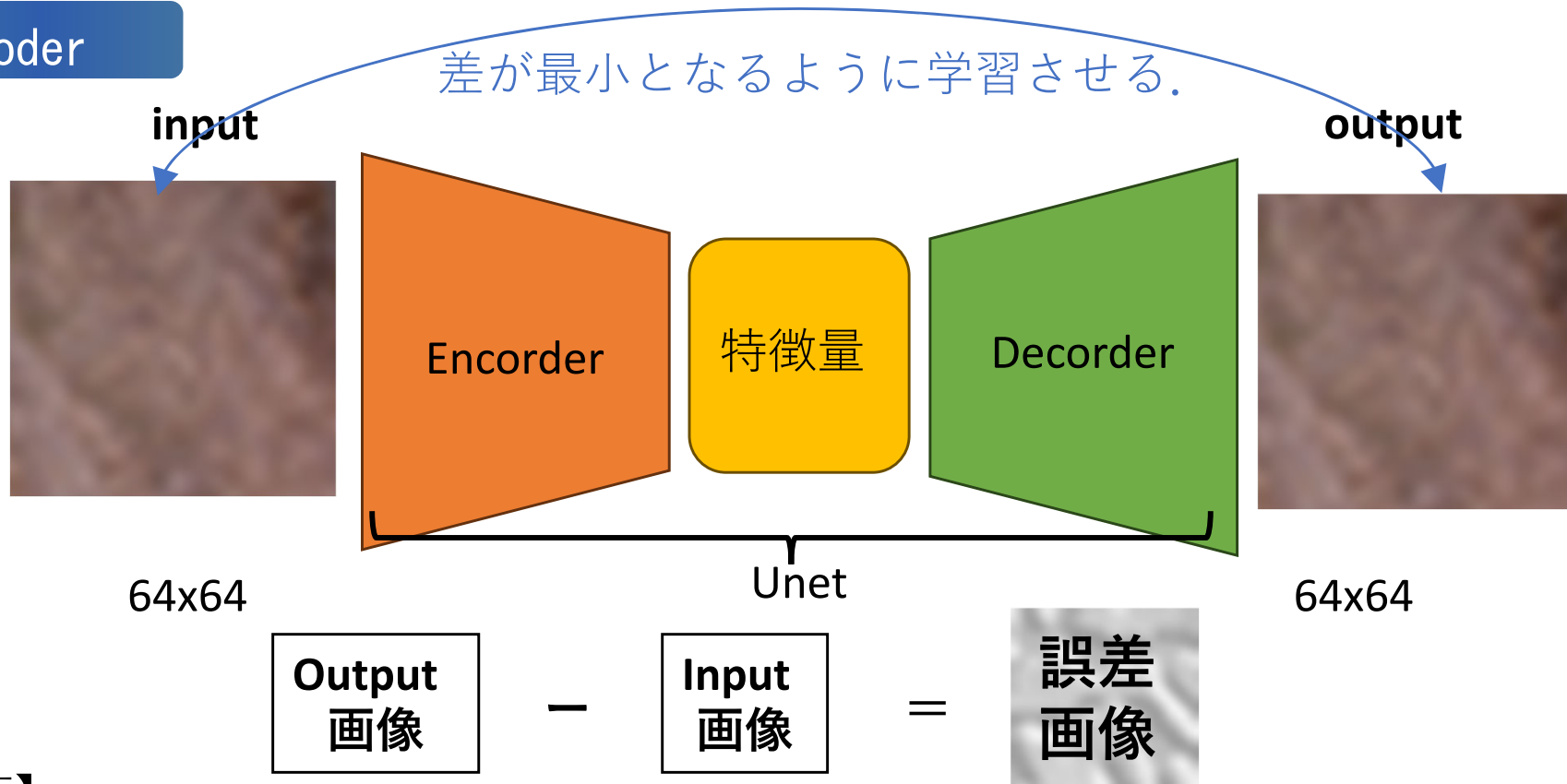
- ・ VLMによって精度が大きく違う
- ・ Qwenは10mで高い精度で検出。
- ・ ゴミの位置・サイズが正しく認識できていない。

		草地									階段								
		10m			30m			60m			10m			30m			60m		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Gemma3	4b	0.025	0.064	0.036	0.167	0.152	0.159	0.000	0.000	0.000	0.033	0.070	0.045	0.068	0.058	0.062	0.000	0.000	0.000
	12b	0.038	0.231	0.065	0.044	0.013	0.020	0.333	0.013	0.025	0.043	0.186	0.070	0.000	0.000	0.000	0.000	0.000	0.000
	27b	0.188	0.192	0.190	0.154	0.051	0.076	0.250	0.013	0.025	0.060	0.035	0.044	0.111	0.012	0.021	0.200	0.012	0.023
LLaVA	7b	0.004	0.013	0.006	0.095	0.025	0.040	0.000	0.000	0.000	0.117	0.244	0.159	0.040	0.023	0.029	0.000	0.000	0.000
	13b	0.011	0.039	0.017	0.035	0.038	0.037	0.000	0.000	0.000	0.006	0.012	0.008	0.000	0.000	0.000	0.008	0.012	0.010
	34b	0.137	0.410	0.206	0.029	0.038	0.033	0.000	0.000	0.000	0.005	0.012	0.007	0.000	0.000	0.000	0.011	0.024	0.015
Ministral3	3b	0.116	0.103	0.109	0.222	0.127	0.161	0.333	0.026	0.048	0.077	0.035	0.048	0.200	0.081	0.115	0.143	0.012	0.022
	8b	0.047	0.103	0.064	0.097	0.127	0.110	0.114	0.052	0.071	0.053	0.081	0.064	0.155	0.149	0.152	0.053	0.024	0.033
	14b	0.359	0.244	0.290	0.344	0.139	0.198	0.667	0.026	0.050	0.222	0.093	0.131	0.539	0.081	0.140	0.000	0.000	0.000
Qwen3-vl	2b	0.781	0.321	0.455	0.500	0.013	0.025	0.000	0.000	0.000	0.880	0.256	0.396	0.000	0.000	0.000	0.000	0.000	0.000
	4b	0.833	0.385	0.526	0.333	0.013	0.024	0.000	0.000	0.000	0.769	0.349	0.480	0.667	0.023	0.044	0.000	0.000	0.000
	8b	0.882	0.385	0.536	0.333	0.025	0.047	0.000	0.000	0.000	0.865	0.372	0.520	0.667	0.023	0.044	0.000	0.000	0.000
	32b	0.882	0.385	0.536	0.214	0.038	0.065	0.000	0.000	0.000	0.769	0.698	0.732	0.250	0.012	0.022	0.000	0.000	0.000



AIを用いた検出(3)

Autoencoder



【検出手順】

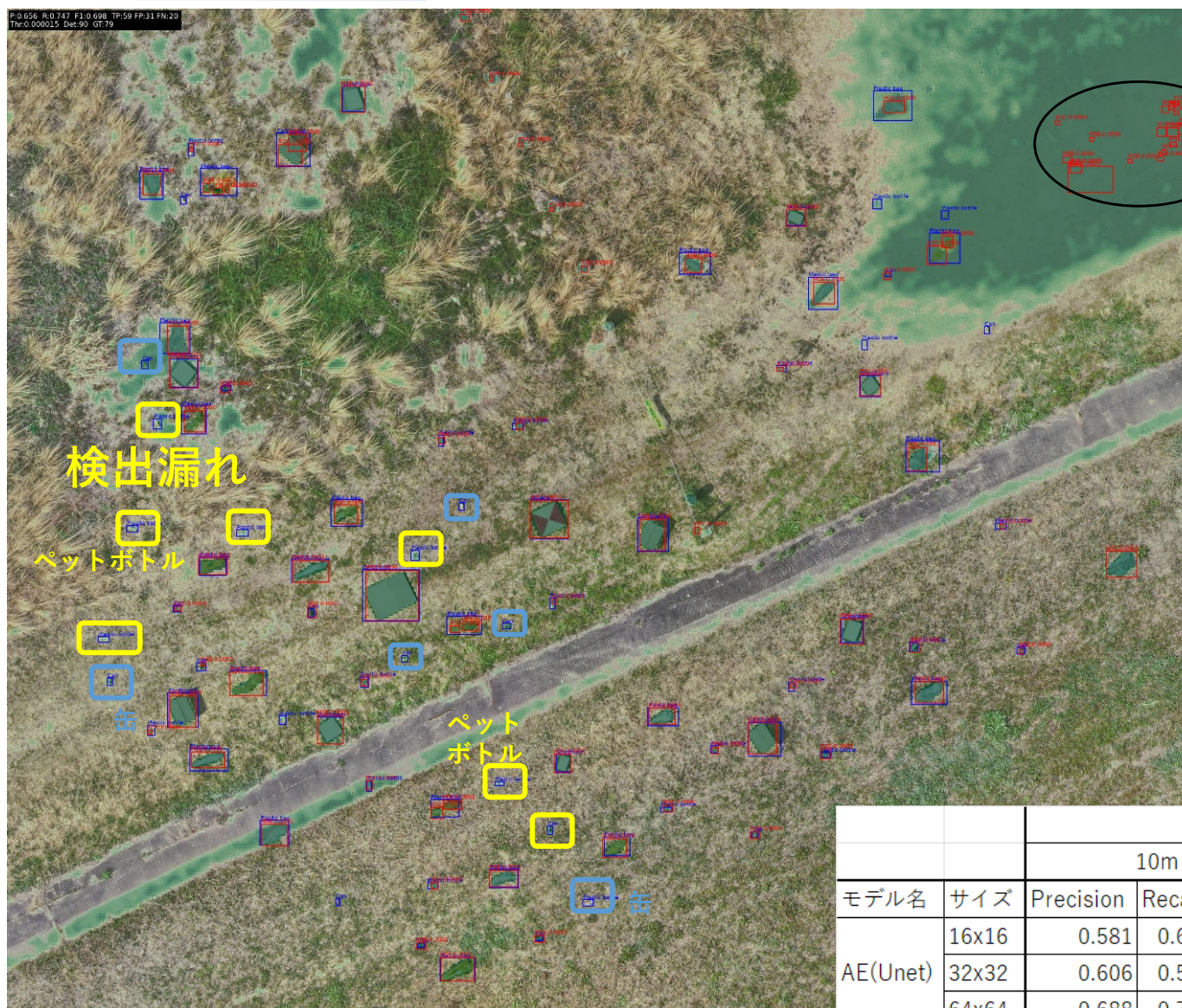
- ①ゴミの無い（正常な）河川敷の画像を学習させる
- ②河川敷の画像をタイルに分割して推論する
ゴミがある場合：ゴミのを学習していないので、復元誤差が大きくなる。
ゴミがない場合：学習しているデータに近いため、復元誤差が小さくなる。
- ③復元誤差の閾値と比べ大きいタイルはゴミが含まれる。＝ゴミの検出
- ④復元誤差のヒートマップを作成し、閾値で包絡する矩形を1つのゴミとする。

平均誤差＝異常度
閾値から異常を検出

AIを用いた検出(3)

2026年8月 DXサマーシンポジウムにて、
アルゴリズムや詳細な検討結果を
ご報告予定です。

Autoencoder



— 正解
— Autoencoder

誤検出が出ている。

裸地で誤検出が出ている
(裸地を学習していない為)

缶・ペットボトルの一部で
検出漏れが出ている。

学習している草地の範囲では、
誤検出が少なく教師なしモデルでも
ゴミ検出が可能

草地:30m タイルサイズ 16x16

		草地								
		10m			30m			60m		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
AE(Unet)	16x16	0.581	0.684	0.628	0.656	0.747	0.698	0.694	0.442	0.540
	32x32	0.606	0.506	0.552	0.842	0.405	0.547	0.879	0.377	0.527
	64x64	0.688	0.705	0.696	0.714	0.506	0.593	0.788	0.338	0.473

高度10mでは物体検出より高精度



検査

高度が上がるにつれて、Precisionが上昇している。 Recallも高い水準。

AEとVLMを用いたゴミ検出手法

- VLM + Autoencoder
- Autoencoder + VLM

先行研究：Object Detection + VLM
物体検出を行い，VLMで精査・ラベル付け



AEとVLMを用いたゴミ検出手法(1)

VLM + Autoencoder

①VLMの検出結果とAutoencoderの復元誤差ヒートマップを用いて、ゴミの検出位置を補正する。

(Bounding Boxの中心をヒートマップのピークに移動)

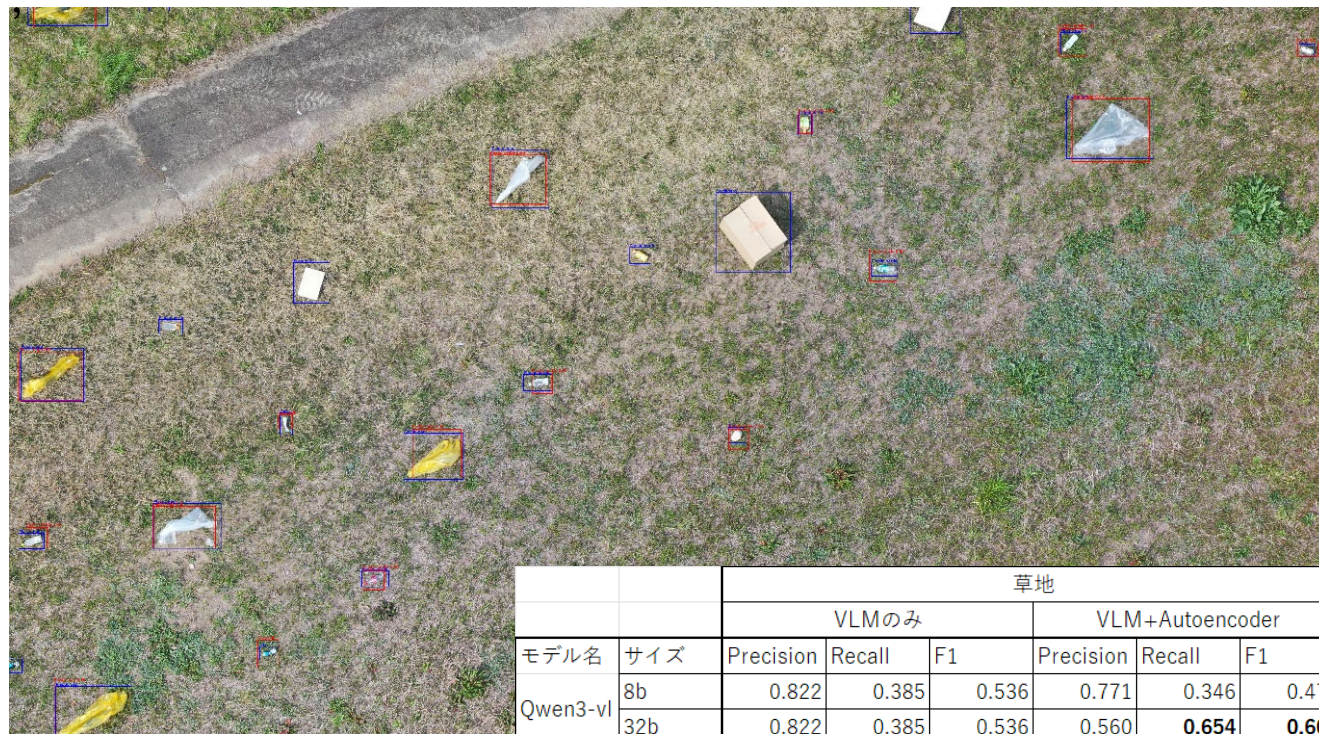
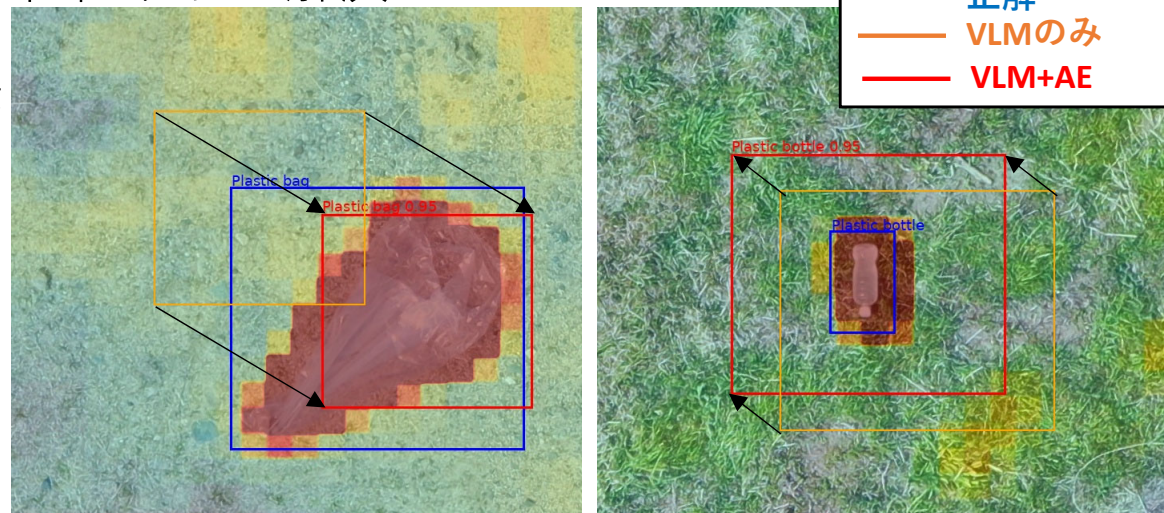
②Autoencoderのゴミ検出アルゴリズム(ヒートマップの閾値による輪郭検出)を用いて、サイズの補正を行う。

- 事前学習済みのVLMを用いる事で様々なゴミを検出する事が可能.
- 位置・サイズのずれはAEを用いて補正が可能.

位置ずれを補正し、Recallが上昇
近くの別のゴミと移動先を誤り、Precisionが低下

VLMのみと比べ、
Recall, F1が大幅に改善

位置ずれの解決



		草地					
		VLMのみ			VLM+Autoencoder		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1
Qwen3-vl	8b	0.822	0.385	0.536	0.771	0.346	0.478
	32b	0.822	0.385	0.536	0.560	0.654	0.604

Autoencoder + VLM

- Autoencoderを持ちいた方法では、復元誤差の閾値設定が重要となる。
閾値を高めに設定すると、検出漏れが増え、低めに検出すると誤検出が増える。

- ①Autoencoderのゴミ判定の閾値を低めに設定(Autoencoder単独での評価の半分, $1.0e-4$)
- ②誤検出含め検出数が増える。(正解も増える)
- ③Autoencoderで検出したゴミをVLM(Qwen3-vl:32b)で再評価し、
背景とした場合はゴミの評価から除外。
Autoencoderでは種類を決定する事ができないが、
VLMを用いる事でラベルを付ける事が可能になる。

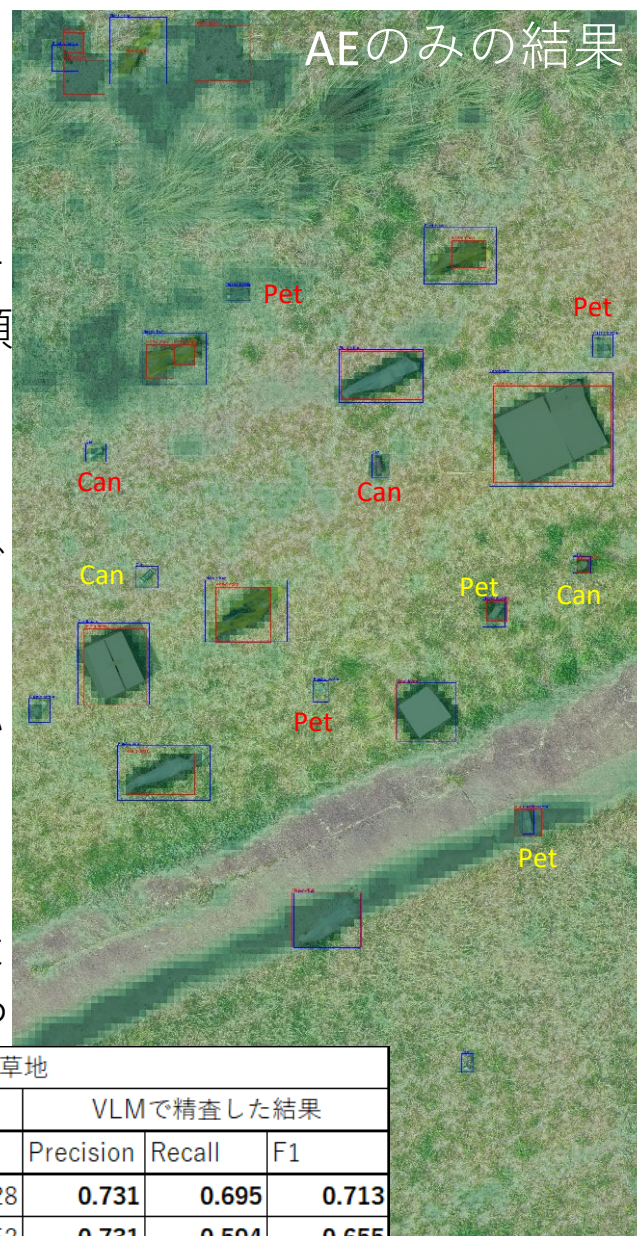
【Autoencoder+VLMへの期待】

➡検出漏れが減り誤検出も減る。 = ゴミ検出精度の向上

AEとVLMを用いたゴミ検出手法(2)

Autoencoder + VLM

- ①AEによる検出結果をVLMでチェックする。
- ②AEではゴミの種類が不明だが、VLMでラベルを付ける事で、ゴミの位置・種類を検出する事が可能。
- ③AEの検出閾値を下げて、検出漏れを減らす。
誤検はVLMで精査する事で減らすことが可能。
×小さいゴミ（缶など）は、VLMが誤ってゴミではないと評価する場合があります、新たな検出漏れとなる場合がある。
- ④VLMによる補正で、草地は全ての指標が改善している



		草地					
		Autoencoderのみ			VLMで精査した結果		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1
Auto encoder	16x16	0.581	0.684	0.628	0.731	0.695	0.713
	32x32	0.606	0.506	0.552	0.731	0.594	0.655
	64x64	0.688	0.705	0.696	0.923	0.758	0.832

←今回検討したモデル内で最高精度



階段への適用

- Autoencoderの適用上の課題
- Autoencoder + VLM

先行研究：Object Detection + VLM



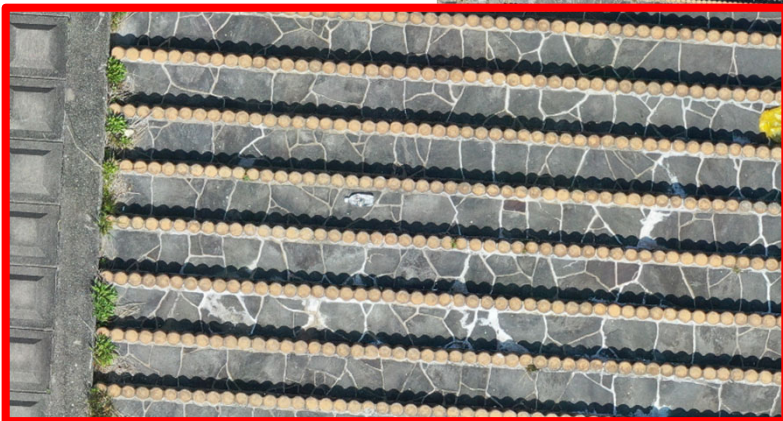
階段への適用

Autoencoder適用上の課題

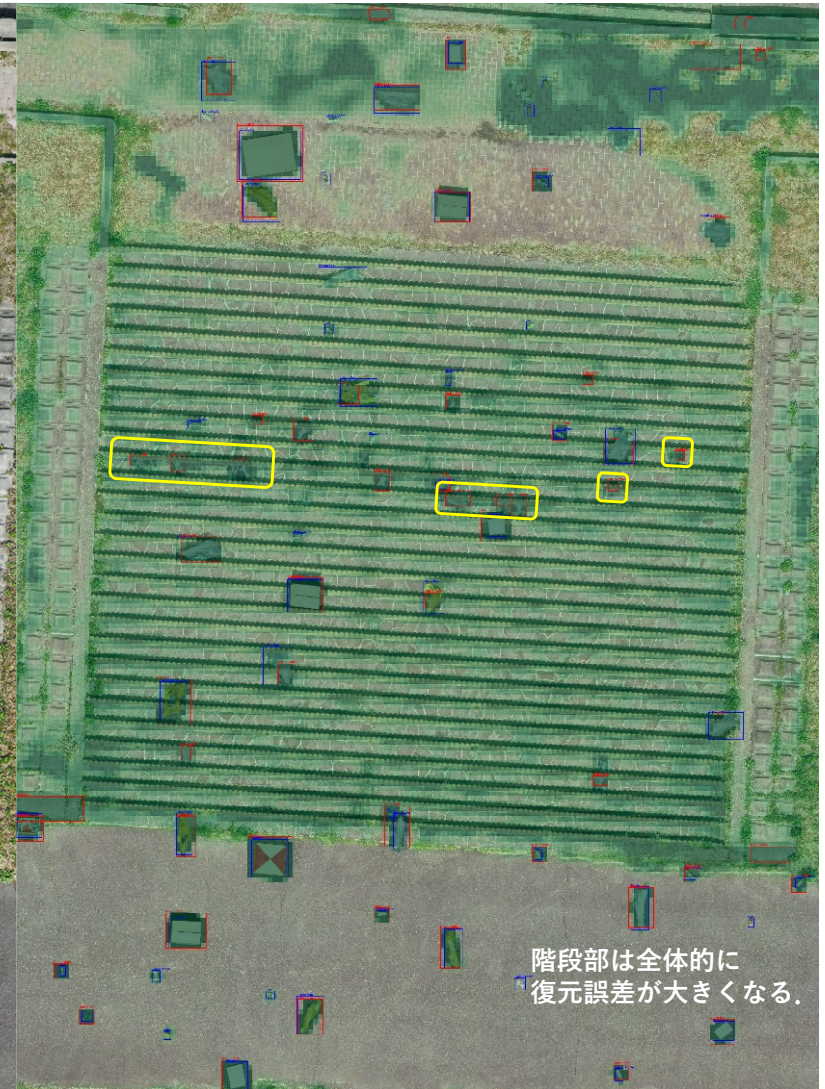
階段の模様を
誤検出している。

不規則な模様を
AEで学習させる際
工夫が必要。
MC-SSIMだけでは、
十分ではない？

階段:10m



Autoencoder(64x64)



階段部は全体的に
復元誤差が大きくなる。

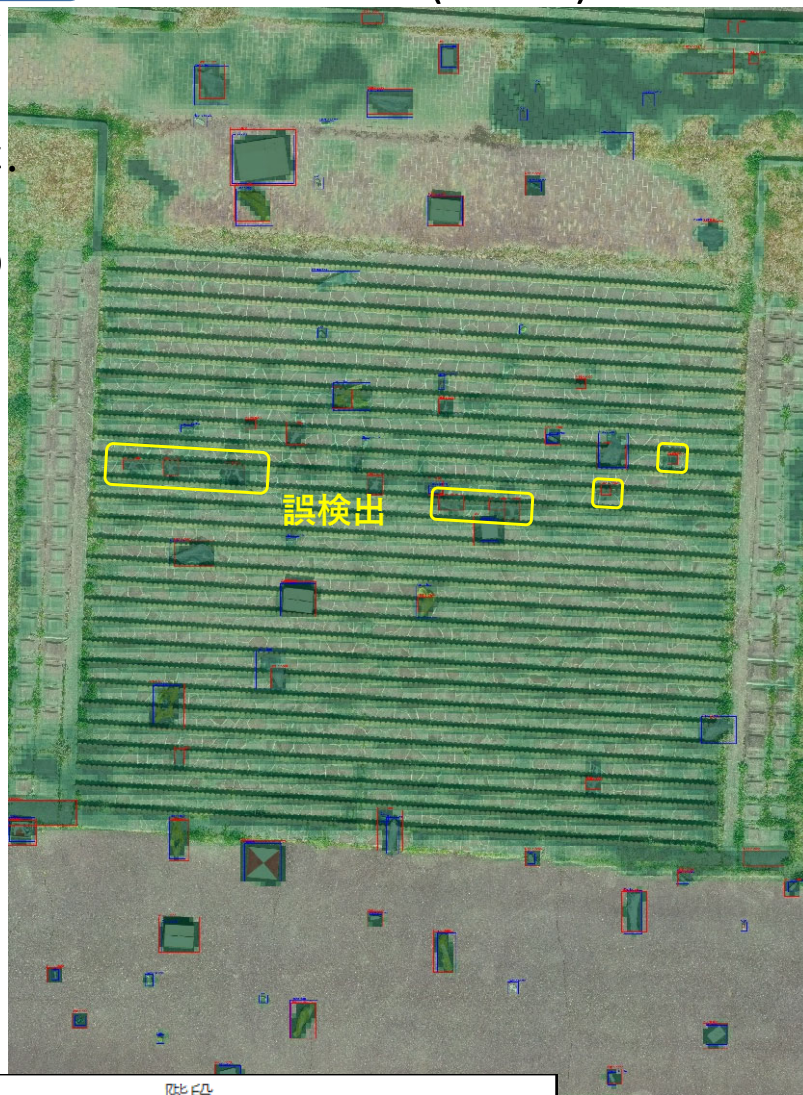
階段部の誤検出を減らす為に、閾値を上げると、
検出漏れが生じやすくなる。

階段への適用

Autoencoder + VLM

- ①AEの結果をVLMで精査する事で、誤検出が低減した。
- ②小さいゴミ（缶・ペットボトル）の検出漏れが一部発生している。

Autoencoder(64x64)



AE+VLM(Qwen3-vl:32b)



		階段					
		Autoencoderのみ			VLMで精査した結果		
モデル名	サイズ	Precision	Recall	F1	Precision	Recall	F1
Auto encoder	16x16	0.449	0.460	0.455	0.930	0.459	0.615
	32x32	0.600	0.414	0.490	0.919	0.456	0.609
	64x64	0.510	0.605	0.553	0.721	0.377	0.495

階段部では、 Precisionが大きく改善した。

草地では最高性能だったが、
階段では32x32が最良となっており、 画像依存性がある。

まとめ



まとめ

推論時間(Nvidia H200)

単位(秒)	適用画像	草本			階段		
	モデル名	10m	30m	60m	10m	30m	60m
Object Detection	Yolo 26n	18	2	1	17	2	1
	Yolo 26x	26	3	1	21	2	1
VLM	Qwen3-vl:8b	1,764	378	348	1,127	664	551
	Qwen3-vl:32b	3,975	1,872	1,415	1,714	976	882
Auto encoder	16x16	780	78	78	609	87	87
	32x32	624	78	78	522	87	87
	64x64	546	78	78	435	87	87
VLM+AE	8b+64x64	2,470	477	380	1,799	773	594
AE+VLM	64x64+32b	763	189	135	588	208	174

高速に処理が可能

10,000x10,000ピクセルの処理に1時間以上かかる

10mの場合は10分程度かかるが、
30m以上では2分以内に処理が終わる。

VLM単体より解析時間が長くなる。

AutoEncoderの処理と同程度の時間で解析可能。

まとめ

- 様々なAIモデルを用いて、仮想的な投棄されたゴミの検出を行った。
- Object Detectionは高速で検出精度も高いが、学習しているデータと解像度が異なる場合に精度が低下した。
- AEやVLMはそれぞれ長所はあるが、誤検出や検出漏れの課題があった。
- AE+VLMの方法では、Precisionが大幅に改善し、階段においても誤検出を抑制し、高い精度でゴミの検出ができる事を確認した。
- AE+VLMは高度10mの画像(10k×10kピクセル)でも、10分程度で解析可能。
- 教師なし+事前学習済みVLMを組合わせたモデルを提案し、予め教師データを用意しなくても、UAV画像から河川ゴミを教師ありモデルと同程度に検出可能である事をしめた。

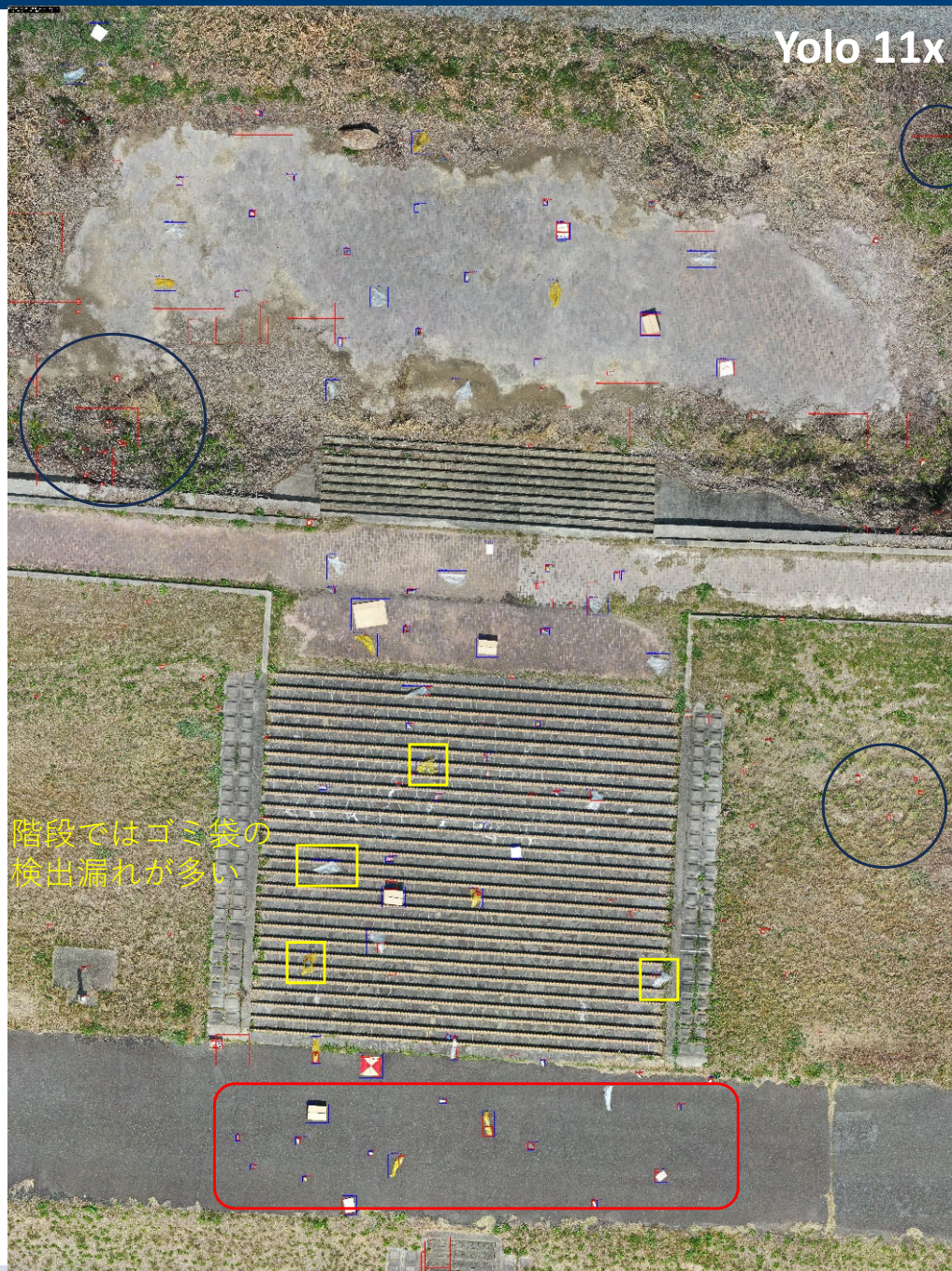
河川技術シンポジウム2026
2026年6月19日（金）

ご清聴ありがとうございました。



株式会社東京建設コンサルタント
TOKEN C. E. E. Consultants Co., Ltd.

草地で誤検出
が出ている



草地で誤検出
が出ている

アスファルト面では缶や
ペットボトルも検出できている。