

(23) 映像から建設機械を検出する AI の汎化性能に関する検討

早川健太郎¹・増田裕正²・蒔苗耕司³

¹正会員 株式会社安藤・間 建設本部技術研究所 (〒305-0822 つくば市荻間 515-1)

E-mail: hayakawa.kentaro@ad-hzm.co.jp

²非会員 富士ソフト株式会社 MS 事業部 (〒101-0022 東京都千代田区神田練堀町 3)

E-mail: masuda@fsi.co.jp

³正会員 宮城大学教授 事業構想学群 (〒981-3298 宮城県黒川郡大和町学苑 1 番地 1)

E-mail: makanae@myu.ac.jp

建設現場を撮影した映像から、施工管理に必要な情報を取得する AI を活用する取り組みが盛んである。筆者らは、映像中の建設機械を自動検出する AI を開発したが、学習データを収集した現場と異なる現場に適用した場合の検出精度低下が課題であった。そこで、機械学習に用いる学習データの構成を工夫することで、別現場へ適用した場合の検出精度を高める検討を進めた。

2 種類の学習データセットを機械学習し、別現場の未知データを入力した結果、建設機械の向きとカメラからの距離に関する学習データをバランスよく含んだ AI で検出精度が高くなった。様々な建設機械の姿を学習したことで、特定の状況に左右されることなく、別現場にも対応する AI になったと考える。また、別現場の映像を学習データとして追加することで、さらに検出精度が向上することを確認した。

Key Words: AI, teaching data, generalization ability, machine learning, object detection

1. はじめに

建設現場を撮影した映像から、安全管理や進捗管理などの施工管理に関する情報の分析や構造物の劣化評価をする取り組みが盛んに行われている¹⁾。近年では AI を用いる手法が注目されているが、AI は学習範囲外の状況に弱く、臨機応変な対応ができないことが課題とされている²⁾。様々な現場状況に対応可能な AI を作成するには、適用する現場ごとに学習データを収集する方法が考えられるが、都度の大量の学習データの収集には時間とコストが掛かることが問題である。筆者らは、機械学習に用いる学習データの構成を工夫することでこれらの課題を解決可能であると考え、学習データセットと現場適用に関する検討を進めたので、本稿にて紹介する。

2. 建設機械の自動検出システムの概要

AI による建設機械の検出精度を評価するため、筆者らがこれまでに開発した映像中の建設機械を自動的に検

出するシステム³⁾ (以下、建機検出 AI と呼称する) を用いた。建機検出 AI は、盛土施工で一般的に使用されている 4 種類の建設機械 (ダンプ、バックホウ、ブルドーザ、振動ローラ) を検出対象としており、入力された未知の画像から 4 種類の建設機械を検出することができる。建機検出 AI の運用フローを図-1 に示す。

- ① 現場に設置した定点カメラで建設機械の作業状況を撮影し、得られた映像から画像データを切り出す。
- ② 画像データに映る建設機械にラベル付け作業を実施し、建設機械の外観と名称を関連付けた学習データセットを作成する。
- ③ 学習データセットを Microsoft 社 Azure Cognitive Services の Custom Vision に入力し、機械学習を行うことで建機検出 AI を作成する。

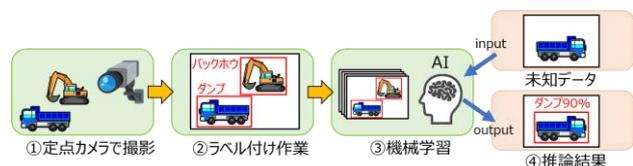


図-1 建機検出 AI の運用フロー

- ④ 建機検出 AI に未知の画像データを入力すると AI による推論が行われ、その画像中に存在する建設機械を検出し、図-2 のような結果が出力される。

3. 2種類の学習データセットと適用現場

図-3 に示すフローで、状況が異なる現場に対する建機検出 AI の検出精度を確認した。現場 A の映像から作成した構成の異なる 2 種類の学習データセットを学習し、2 つの AI (無作為 AI, バランス AI と呼称) を作成した。

2 種類の AI の特徴を表-1 に示す。現場 A の映像のうち、建設機械が稼働している時間帯から無作為に 1,800 枚の画像を切り出し、9,679 個のラベルを付与したデータセットを学習したものを「無作為 AI」とした。もう一方の学習データセットでは、建設機械の向き (図-4) とカメラからの距離 (図-5) に注目して現場 A の映像から画像を切り出した。ラベルを付与する際に建設機械の

名称に加えて向きと距離に関する情報も与えており、例として【ダンプ、左向き、近距離】のようなラベルとなる。学習データセットに含まれる建設機械の外観が特定の状況に偏らないように、向きに関する 8 種類のラベル数と距離に関する 3 種類のラベル数がそれぞれ均一になるように意図的にバランスを取り、5,880 個のラベルを採用した。この学習データセットを学習したものを「バランス AI」とした。図-6, 7 は、無作為 AI とバランス AI の学習データセットに含まれるダンプ画像のうち、向きに関するラベル数を表したものである。無作為 AI はそれぞれの向きでラベル数に差異があるが、バランス AI は均一になっている。

2 種類の AI に入力する未知データは現場 A (図-8) と現場 B (図-9) の映像から抽出した。現場 A の未知データは、学習データ収集時と同一の定点カメラから得てお



図-2 出力結果

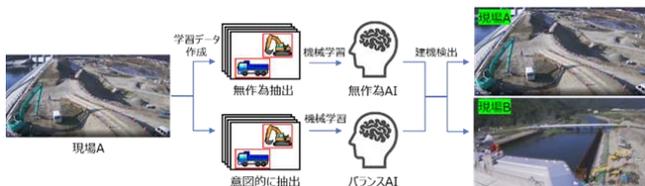


図-3 学習させた現場と異なる現場への適用フロー

表-1 無作為 AI とバランス AI の特徴

	学習データを収集した現場	ラベル数	画像数	ラベル数のバランス考慮
無作為 AI	現場 A	9,679	1,800	なし
バランス AI	現場 A	5,880	2,623	向き、距離



図-4 建設機械の向き



図-5 建設機械の距離

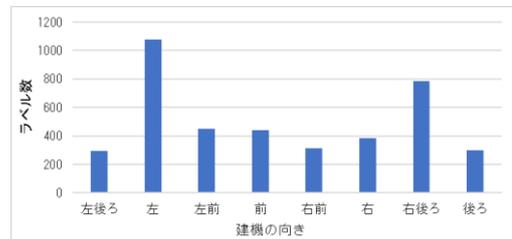


図-6 無作為 AI のラベル数 (ダンプ向き)

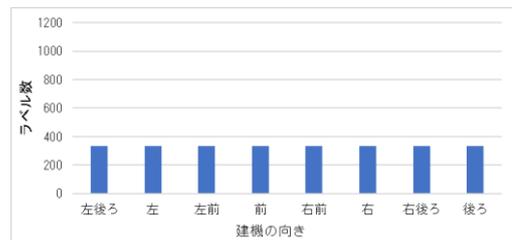


図-7 バランス AI のラベル数 (ダンプ向き)



図-8 現場 A の映像



図-9 現場 B の映像

り、カメラから約 100m 以内の近～中距離で防潮堤の盛土作業を行う建設機械が多いことが特徴である。現場 B の場合、カメラの型式は同一だが、カメラから 100～150m の遠距離で河川堤防の構築を行う建設機械が多く、カメラアングルや背景の状況も現場 A と異なる。未知データを用いた AI の検出精度評価指標として IoU (Intersection over Union) と Accuracy を用いた。IoU は図-10 に示すように、目視で与えられた正解ラベルの領域と、AI による検出領域の重なりが一定値を超えたとき、正しく検出できたと判定する手法である。ここでは IoU が 30% 以上になったとき、正解と見なしている。Accuracy は AI による検出結果の正解率を表す指標の 1 つであり、 $Accuracy = \frac{\text{正解数}}{\text{正解数} + \text{誤検出数} + \text{未検出数}}$ で求められる。

4. 適用結果

現場 A と現場 B の未知データを無作為 AI とバランス AI に入力し、正解率を求めた結果を表-2 に示す。現場 A に適用した場合、無作為 AI の正解率は 63.23%、バランス AI が 57.97% となった。一方、現場 B では、無作為 AI は 49.43% まで正解率が低下したのに対し、バランス AI は 61.17% となり、異なる現場に適用しても正解率が向上する結果となった。

この結果は、学習データセットの作成方法において、その現場の特徴が AI に色濃く反映されたためであると考える。すなわち、現場 A の映像中の建設機械は、カメラから近～中距離で作業することが多く、搬入経路の都合上、特定の向きの建設機械が頻出している。このような偏った特徴を持つ学習データセットを採用した無作為 AI は、同様な偏りを有する現場の建設機械を検出することに特化し、状況が異なる現場の未知データに対する

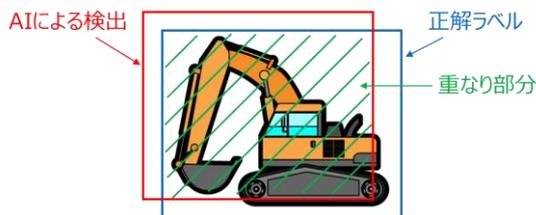


図-10 IoU を用いた判定方法

表-2 2つの AI の正解率

	Accuracy (%)	
	現場A	現場B
無作為AI	63.23	49.43
バランスAI	57.97	61.17

正解率が低下する。一方、建設機械の向きと距離のラベル数をバランスよく整えた学習データセットを採用したバランス AI は無作為 AI と比較して、特定の状況に左右されにくく、異なる現場にも対応する汎化性能の高い AI になったと考えられる。

5. 学習データの追加による検出精度の変化

現場 B に無作為 AI とバランス AI を適用した場合の正解率を向上させるため、現場 B で新たに取得した学習データを AI に追加して学習し、正解率の変化を検証した。この AI の特徴を表-3 に示す。3 章で述べた無作為 AI とバランス AI の学習に用いた学習データセットに加えて、現場 B の映像から画像を切りだし学習データとして追加している。追加した学習データは建設機械が稼働している時間帯から無作為に画像を切り出し、2,914 個のラベルを付与した。また、比較用として、現場 B で新たに取得した学習データのみで AI を作成した。

学習データを追加した AI に対して、現場 B の未知データ (3 章と同一) を入力し、正解率を求めた結果を表-4 に示す。無作為 AI に学習データを追加したものでは、正解率が 63.11%、バランス AI に学習データを追加したものでは正解率が 71.00% と双方が学習データの追加によって正解率が向上している。また、追加用の学習データのみで作成した AI は 54.17% となった。

この結果から、現場 B の未知データから建設機械を検出するとき、現場 A の画像のみを学習していた 2 つの AI に対して、現場 B の画像を追加で学習することは有効であることが確認された。特にバランス AI に学習データ

表-3 現場 B の学習データを追加した AI の特徴

	学習データを収集した現場	ラベル数	画像数	ラベル数のバランス考慮
無作為AI	現場A	9,679	1,800	なし
	現場B (追加データ)	2,914	1,020	なし
バランスAI	現場A	5,880	2,623	向き、距離
	現場B (追加データ)	2,914	1,020	なし

表-4 学習データを追加した AI の正解率

	Accuracy (%)	
	学習データ追加前	学習データ追加後
無作為AI	49.43	63.11
バランスAI	61.17	71.00
追加の学習データのみ	-	54.17

を追加したものでは、現場 B のラベル数が 2914 と比較的小数であるにもかかわらず、正解率が 70%以上と高い精度となっている。無作為 AI も学習データ追加後に正解率が向上しているが、現場 A の偏った特徴を持つデータセットを採用して特定の状況に特化しているため、バランス AI ほどの正解率には至らなかったと考えられる。また、追加用の学習データのみから作成した AI は現場 B の状況に特化しているため高い正解率を持つと思われたが、絶対的なラベル数が不足していることにより約 54%の正解率となった。

6. 建機検出 AI の運用方法の提案

建機検出 AI をより効果的に運用するためのフローを図-11 に示す。従来、建機検出 AI の運用においては、現場ごとに大量の学習データを収集する時間とコストが課題であったが、学習データセットの構成を工夫し、異なる現場でも正解率を保てるような汎化性能の高いバランス AI を用いることで解決が可能である。建設機械の向きと距離のラベル数を均一化した学習データの作成には通常よりも時間を要するが、適用対象とする現場とは異なる現場で事前に学習データセットを作成しバランス AI を作成するため、対象現場へスピーディに適用できる。現場に適用後、より高い正解率を求める場合、対象現場の映像から追加の学習データを収集し、バランス AI に学習させることで精度が向上する。追加の学習データは建設機械の向きと距離のラベル数を考慮しない無

作為なものでよいため、短時間のうちに学習データを収集できる利点がある。

以上のように汎化性能の高い建機検出 AI を運用することで、現場導入の都度の学習データ収集に要する時間とコストを大幅に効率化できると考える。

7. 今後の展望

汎化性能の高い建機検出 AI により、多数の現場への展開が容易になる一方で、事前作業であるラベルの選別作業と均一化には多大な労力を要し、無作為なラベル付与作業よりも時間が掛かる。今後はラベルの付与、選別を効率化する手法を検討していきたい。

参考文献

- 1) Kareem Mostafa, Tarek Hegazy : Review of image-based analysis and applications in construction, Automation in Construction, Vol.122, 103516, 2021.
- 2) 全 邦 釘 : 土木工学分野における人工知能技術活用のために解決すべき課題と進めるべき研究開発, 第 1 回 AI・データサイエンスシンポジウム論文集, pp.9-16, 2020.
- 3) 早川健太郎, 黒台昌弘, 増田裕正, 蒔苗耕司 : AI による建設機械検出システムの開発と検出精度を向上する試み, 第 1 回 AI・データサイエンスシンポジウム論文集, pp.313-319, 2020.

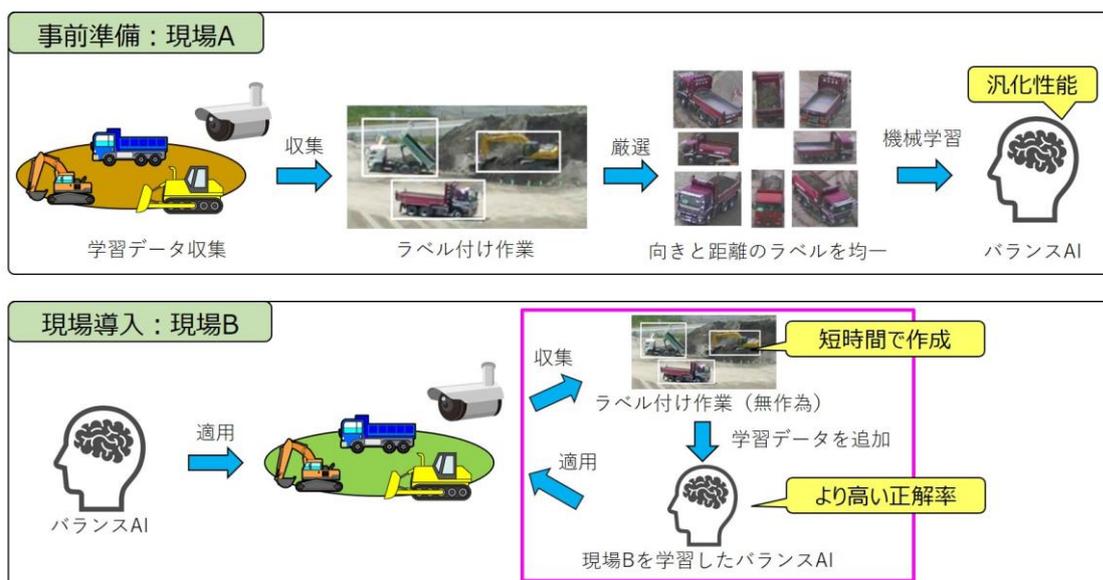


図-11 建機検出 AI を効果的に運用するフロー