

AI を用いて映像中の建設機械を自動的に検出するシステムの開発

安藤ハザマ
富士ソフト

○早川健太郎 木付拓磨
増田裕正 寺原 勲

1. はじめに

建設現場では、レーザスキャナや GNSS 等の 3 次元計測機器を活用した ICT 施工が盛んに行われている。様々な機器を用いて現場の形状を計測することで出来高や出来形を管理しているが、これらの多くは施工完了時の一瞬を計測したものであり、あくまで施工結果の確認として実施されている。結果に至るまでの歩掛や日々の出来高等のプロセス管理は異なった手法で行われることが多く、定点カメラで撮影した映像を用いて施工管理を行うものなどが挙げられる。例えば、動画やタイムラプスカメラで工事の進捗を一定間隔で記録するものがある。これらの活用方法では、毎秒数フレーム～数分間隔の高い時間分解能で現場内を継続的に記録することが可能であり、映像から施工の結果だけでなくプロセスを確認することができる。しかし、映像を見るだけでは、工事完成形に対する進捗状況が直観的にわかりにくい、施工量や距離、面積等の定量的な情報を取得しづらい、稼働している建設機械の台数等の常に変化する情報を素早く把握できないといった課題がある。そこで、上記の課題を解決しつつ建設現場の進捗管理を効率的に行うため、「4K 定点カメラ映像による工事進捗管理システム」¹⁾²⁾を開発した。このシステムは①3D

データの重畳表示、②映像から距離や面積を算出、③AIによる進捗レポート、④オルソ画像の作成の4つの機能を搭載している。

このうち③AIによる進捗レポートは、AIを用いて映像中の建設機械を検出し出来高管理に活用するものであり、特に先進的な機能である。従来は職員が現場に長時間留まり、建設機械の歩掛や稼働状況を知ることによって出来高を管理していたが、AIによる建設機械の検出によりこの業務を効率化することが可能である。本稿では「AIを用いて映像中の建設機械を自動的に検出するシステム」について詳しく述べる。

なお、本件は国土交通省「建設現場の生産性を飛躍的に向上するための革新的技術の導入・活用に関するプロジェクト」(PRISM)に採択され開発したシステムの一部である。

2. 建設機械検出システムの開発

2.1 システムの概要

映像中の建設機械を自動的に検出するシステムの概要を図-1に示す。このシステムの特徴は、建機検出AIとカテゴリ認識AIと呼称する2つのAIで構成されていることである。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

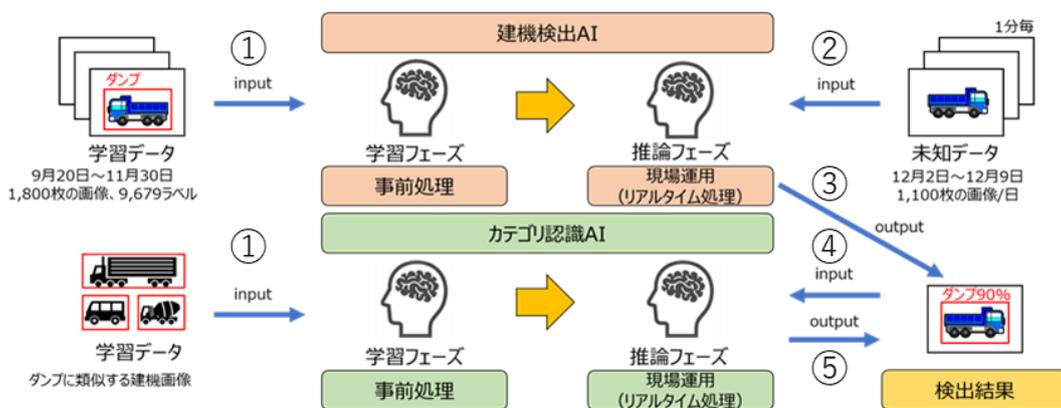


図-1 建設機械検出システムの概要

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。この検出結果では外観がダンプに類似する車両を誤検出する傾向があるため, それらを区別することが可能なカテゴリ認識AIを用いた。建機検出AIがダンプを検出した場合, それが真にダンプであるかさらに分類することで, 検出精度を高めている。図中の①は機械学習によるAIの生成を表しており, システム運用前に実施する。ここでは事前に収集した複数のデータを学習することで, 建機検出AIとカテゴリ認識AIを作成している。②~⑤は未知の映像から建設機械を検出する手順を表している。②では建設機械を検出したい画像を建機検出AIに入力し, ③で検出結果を出力している。ここでダンプとして検出された場合は外観が類似する車両である可能性があるため, ③で得られたダンプの画像を④でカテゴリ認識AIに入力し, 検出精度を高めた結果を⑤で出力する。これにより未知の映像に存在する建設機械が検出される。

建設現場の映像をこのシステムに入力することで, 映像中の建設機械を自動的に検出することが可能になり, 職員が現場を常時観察せずとも出来高管理の基礎となるデータを得ることができる。

2.2 システム運用までの準備

以下にシステムを現場で運用するために必要な準備事項を示す。

(1) 対象現場と定点カメラの設置

映像撮影の対象として, 図-2に示す延長300m, T.P.+14.5mの防潮堤を構築する工事現場を選定した。この防潮堤に近接する位置に構築された2つの水門上に定点カメラ(JVCケンウッド社製VN-178WPR)を設置し, 4K解像度(3,840×2,160), 1fps間隔で防潮堤を構築する様子を記録した。図-3は定点カメラで実際に得られた画像である。

(2) 検出の対象とする建設機械の選定

盛土施工で一般的に使用されている4種類(ダ

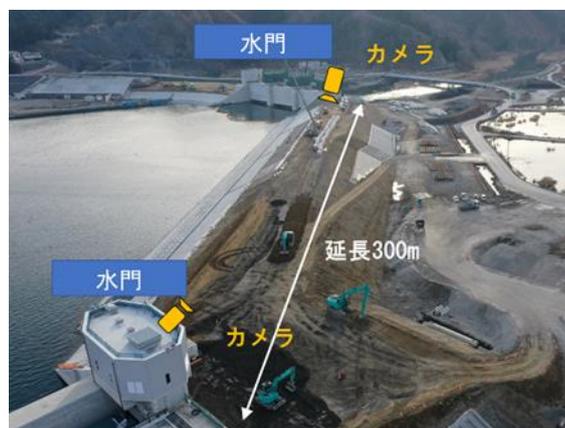


図-2 対象とした防潮堤の工事現場

ンプ, バックホウ, ブルドーザ, 振動ローラ)を対象にした。同種の機械であっても仕様(バックホウのバケット容量やアーム長など)や色が異なる機械が複数台存在するが, ここでは同種の機械として扱っている。また, 検出の対象とするエリアは定点カメラからの距離が150m以内の盛土施工範囲(図-3)とした。

(3) AIの作成に利用するプラットフォームの選定

AIを作成するために, Microsoft社のCustom Visionを利用した。このプラットフォームでは機械学習技術を用いて, 学習データとして入力した画像から学習モデルを作成し, 学習した対象を検出するAIを作成することが可能である。このAIは学習によって得られた対象の特徴を用いて画像内に存在する対象を検出し, 名称と画像内の座標を得ることができる。

(4) 学習用の画像データの取得

盛土施工用の建設機械が稼働していた2019年9月20日~11月30日までの定点カメラ映像のうち, 作業が行われている時間帯から無作為に1,800枚の静止画(JPEG)を切り出し学習用の画像データとした。

(5) 画像データへのラベル付け

学習用の画像データに映っている建設機械の位置と名称を登録する作業をラベル付けと呼び, ラベルを付けた画像データはAIにとっての正解値である。図-4のように建設機械を矩形で囲み位置を決定し名称を登録する作業により, 建設機械の外観と名称が関連付けられ学習データとなる。ラベル付けは通常, 目視かつ手作業で行うため膨大な時間を要する。1,800枚の画像データには総数10,000以上の建設機械が映っており, そのうち9,679台の建設機械にラベル付けを行った。図-5はラベル付けを行った画像データの一例であり, 一枚の画像中に複数台の建設機械が映っている場合は, 全てにラベルを付与している。4種の建設機械のラベル数の内訳は, ダンプ4,043個, バックホウ4,042個, ブルドーザ812個, 振動ローラ782個である。ただし, 矩形の大きさが100×100px以下となる建設機械はカメラからの距離が概ね150mを



図-3 定点カメラから得られる映像

超えるものであり、検出の対象範囲外となるので、ラベル付け作業の対象外とした。

(6) 建機検出 AI の作成

ラベル付けされた画像データを学習データとして、Custom Vision にて機械学習を実行した。これにより、未知の映像中に存在する 4 種の建設機械を検出できる建機検出 AI を作成した。建機検出 AI に未知の映像を入力し、建設機械を検出した結果を図-6 に示す。建機検出 AI は入力された映像内で学習した特徴に似ている対象を検出すると、対象の名称とそれを囲む矩形座標を出力する。システム上では建設機械の種類ごとに異なる色の矩形で表現しており、同時に矩形上部には推論確率(推論の確からしさ)が表示される。図-6 の左上の例では赤い矩形で囲まれた対象は「90.5%の確率でブルドーザである」ことを示している。

(7) カテゴリ認識 AI の作成

建機検出 AI を用いて未知の映像中からダンプ

を検出する際、ダンプと外観が類似している建設機械を誤ってダンプとして検出する事例が多く発生した。図-7 はダンプとして誤検出された建設機械の例である。矩形で囲まれている対象は、正しくは生コン車や資材搬入用のトレーラである。

これらの誤検出結果を正しい結果へ導くため、ダンプとダンプに類似した 5 種の車両のカテゴリを分類するカテゴリ認識 AI を作成し、建機検出 AI による検出結果に対して適用した。AI の作成に Microsoft 社の Custom Vision を利用した点と学習用の画像データの収集期間は建機検出 AI と同様である。学習用の画像データに映っているダンプに類似した車両を 5 種類選択しラベル付け作業を行い、ラベル総数は 248 個(表-1)とした。

建機検出 AI とカテゴリ認識 AI で大きく異なる点は、建設機械の見つけ方にある。建機検出 AI は入力画像の全範囲から対象となる物体の存在を検出している。一方、カテゴリ認識 AI は建機検出 AI の検出結果(矩形で囲まれた画像)のうち、誤判定の多いダンプに対して働くものである。対象は既にダンプとして判定された画像のみであり、それをダンプもしくはダンプ以外の車両に分類してい

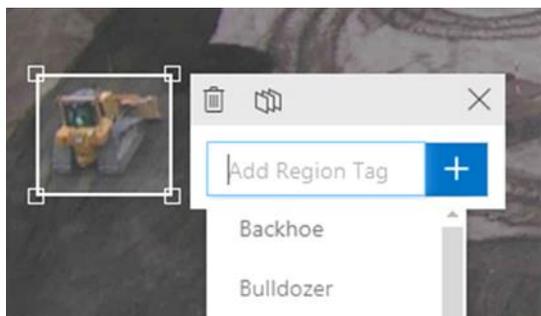


図-4 ラベル付け作業の様子



図-7 建機検出 AI がダンプと誤検出した例



図-5 ラベル付けされた画像データの例

表-1 カテゴリ認識 AI の作成に用いた車両のラベル数

カテゴリ	ラベル数	ラベル合計
ダンプ	50	248
Wキャブトラック	50	
生コン車	50	
トレーラ	48	
乗用車	50	



図-6 建機検出 AI での検出結果

る。この2つのAIを連続的に運用することで、盛土上で稼働する建設機械を精度よく検出することが可能になる。

3. 建設機械検出システムの性能確認と運用結果

3.1 AIの性能確認の方法

一般的にAIの性能を評価する場合、性能指標として「Precision」、「Recall」、「F値」、「mAP」の4つが用いられるが、このうちPrecisionとRecallに注目し性能を評価する。本稿の建設機械検出の事例で言えば図-8に示すとおりである。PrecisionはAIがダンプとして検出した対象のうち、真にダンプだった対象の割合を示している。Precisionの値が高い場合、誤った検出が少ないことを意味している。Recallは真にダンプである対象のうち、AIがダンプとして検出した対象の割合を表しており、Recallの値が高いほど対象の見逃しが少ないことを意味している。

3.2 評価用データ

AIの性能指標を求めるため、建機検出AIとカテゴリ認識AIに評価用のデータを入力して得られた検出結果を評価用データの正解値と比較する。評価用データとして2.2.(4)と同期間に取得した映像のうち、学習データとして使用していない画像から72枚抽出した。その画像に含まれる4種の建設機械を目視し、正解となるラベルを付与した。その内訳はダンプ190個、バックホウ132個、ブルドーザ30個、振動ローラ27個の合計379個である。

3.3 学習用のラベル数と性能指標の関係

一般には、AIを用いて対象を検出する場合、学習データの数が多の方が検出精度は向上すると言われている³⁾。しかし、多くの学習データを用意しラベルを付ける作業には多くの時間を要するため、必要となる学習用ラベルは極力少ない数で高い性能を出すことを考えたい。そこで、建機検出AIのラベル数を増加させた場合の性能指標への影響を検証した。

図-9, 10には評価用データを用いて算出したAI

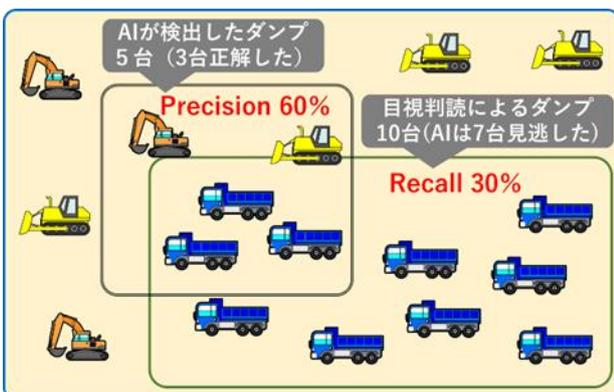


図-8 建機検出AIにおける性能指標

の性能指標を示す。図-9はブルドーザを検出した場合のラベル数と性能指標との関係である。Precisionはラベル数が約400個の時に既に80%以上と高い値を示し、500個を超えると100%となる。すなわち、AIがブルドーザとして検出した対象は、それが全て正解であったことを示している。一方、Recallはラベル数500個までは低い値を示しているが、600個を超えると80%以上となる。図-10にはダンプについての同様のグラフを示す。ダンプは複数台が同時に映像に映る機会が多いため、最小ラベル数が2,424個とブルドーザと比較して多くなっている。そのため、PrecisionとRecallは初期から概ね80~90%と高い水準を維持している。

図-9, 10より、学習データに用いるラベル数が増加するにつれてAIの性能指標が向上していることが明らかとなった。また、ラベル数が一定数を超えると性能指標は収束する傾向にあり、高い値で安定することも確認された。

3.4 AIによる検出結果と目視計数の比較

AIの性能指標が80%を超えたことが確認できたため、現場運用を開始し未知の映像から建設機械の検出を行った。このときの検出誤差を評価するため、AIによる検出回数と目視で計数された台数を比較した。ここでの検出誤差は、(AI検出回数 - 目視台数) / 目視台数で定義した。評価に用いるデータは、AIの学習に使用していない2019年12月2日~12月9日の映像を1分間隔で静止画として切り出したもの(約1,100枚/日)を採用した。建設機械ごとの検出誤差を図-11に示す。目視での計

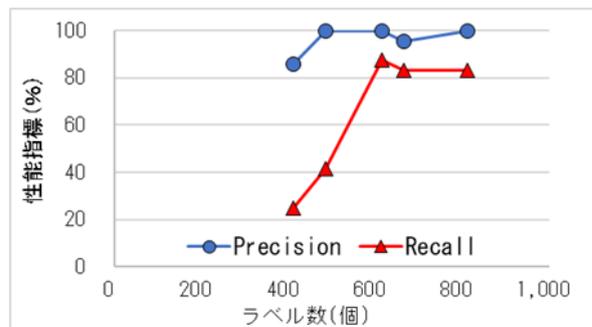


図-9 建機検出AIの性能指標(ブルドーザ)

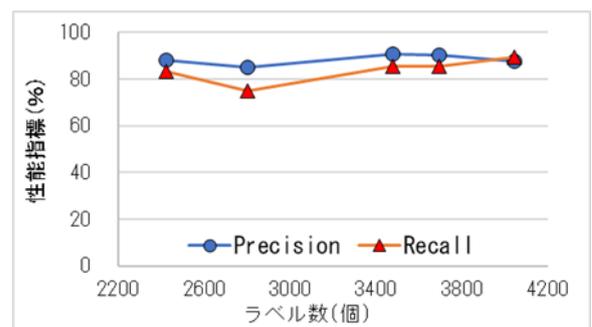


図-10 建機検出AIの性能指標(ダンプ)

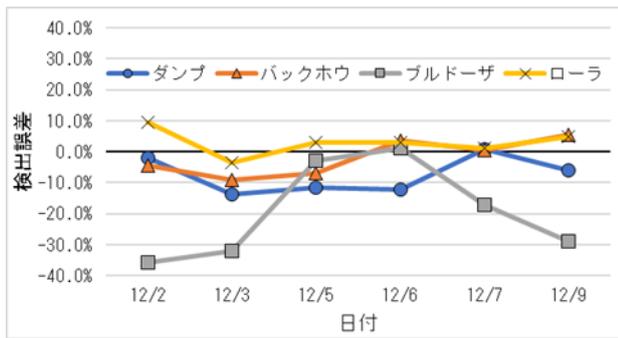


図-11 建設機械ごとの検出誤差

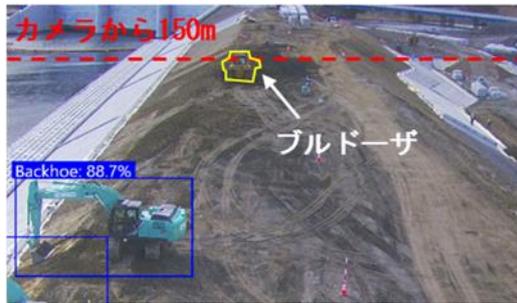


図-12 対象範囲の境界付近のブルドーザ

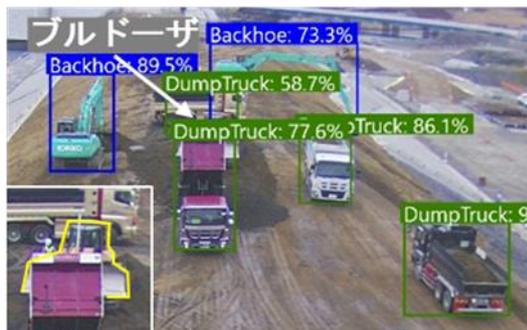


図-13 ダンプと重なった状態のブルドーザ

数に対する AI による検出誤差は、バックホウと振動ローラの場合でいずれの日も $\pm 10\%$ 以内に収まっている。一方、ダンプは $+0.8\% \sim -13.7\%$ 、ブルドーザは $+1.2\% \sim -35.6\%$ の範囲と変動幅が広く、特にブルドーザの検出誤差が大きい。この原因として、次の 2 つが考えられる。第一に、ブルドーザの検出誤差が大きい日の敷き均し作業は図-12 に示すように検出対象範囲とした 150m の境界付近で行っており、ブルドーザは他の建設機械と比べて機体の実寸が小さいため、映像中の解像度が低下したことが挙げられる。第二は盛土材を運搬しているダンプと映像上で重なる(図-13)ことで、建機検出 AI による検出の困難さが増したものと想定される。

このように、AI の性能指標のみで評価した場合は Precision と Recall 共に 80% 以上という結果が出たが、実際に現場で運用し目視結果と比較することで解像度の低下や障害物との重なりを苦手と



図-14 AI が検出したダンプ台数の積み上げグラフ

していることがわかり、性能指標だけでは把握できない課題が判明した。AI の性能を向上させるだけでなく、カメラの設置位置や建設機械の動線などを工夫し図-12, 13 のような状況を避けることで、その性能を発揮することができる状況を整えることも重要である。

4. 出来高管理の高度化

検出対象の解像度の低下や障害物との重なりなどの不適切な条件下を除けば、映像中の建設機械を $\pm 14\%$ 程度の誤差で検出できることが明らかとなり、現場の出来高管理に活用可能であると考えられる。例えばダンプの台数を計数することで、日々の運搬土量を求めることができる。さらに図-14 のようにダンプの延べ台数の積み上げ推移を見ることで、ダンプが集中するもしくは途切れる時間帯を把握し、現場の改善に資することも可能である。

従来、これらの情報は職員による観察で得ていたが、現場に滞在できる時間には限りがあった。このシステムを用いることで終日の情報を自動的に得ることができるため、職員の負担が減るだけではなく、建設機械の時間ごとの推移などこれまでは気づきにくかった情報を得られるようになり、建設現場の管理業務をより効率的に進めることができるようになる。

5. システム運用上の留意点

建設機械検出システムを運用していく中でいくつかの課題が明らかになったので、以下に述べる。

(1) 現場の広さとカメラの解像度

定点カメラを設置して現場の映像を得るとき、対象とする範囲に対して適切な解像度のカメラを選択する必要がある。解像度の低いカメラを使用すると、撮影距離によっては映像に映る建設機械の解像度が低下し、モザイク処理を施したような見た目になる。このような映像は学習データとして不適切であり、検出精度の低下を招く恐れがある。

(2) カメラの設置位置

定点カメラを設置する位置によって撮影対象に偏りが生じる可能性がある。例えば人の目線程度の高さにカメラを設置した場合、特定の建設機械

が視界を遮り他の建設機械が映像に映らないことがある。このような状況が続くと学習データの収集に影響があるだけでなく、現場内の建設機械数を正確に検出できないため、出来高管理のデータとして運用することができなくなる。そのため、カメラは現場を俯瞰するような位置に設置することが望ましい。

(3) 学習データの収集期間

建設現場では、施工の進捗によって地形や動線の変化、使用する建設機械の入替えなど、現場の状況は日々更新されている。このような状況の中で短期間のうちに集中的に学習データを収集した場合、特定の状況での検出に特化した AI となる可能性がある。AI に汎用性を持たせるためにも様々な状況の学習データを得ることが望ましい。

(4) ラベル付け作業時間

ラベル付け作業は人の手作業で数千回にわたって行われるため、多くの人工を要する。今回使用した学習データのラベル数は 9,679 個であり、ラベル付与には約 256 時間を要している。人物の顔や一般的な車両などを検出する場合、公開されている学習データセットを活用することでラベル付け作業を省略することができるが、建設機械のような特殊な物はデータセット公開には至っていない。

(5) AI の検出精度の向上手法

本稿では AI の検出精度を向上させる手法として、カテゴリ認識 AI の活用と学習データに使用するラベル数の増加が有効であったことを示している。一方で、筆者らはその他にも精度を向上する手法を検討しており、ラベルの偏りも重要な要素であることを確認している³⁾。偏りとは、建設機械の向きごとのラベル数の多少を意味しており、AI の検出精度に影響を与えることがわかっている。そのため、無作為にラベル数を増やすだけでなく建設機械の向き等に一定のルールを設定して学習データを作成することも検討すべきである。

6. おわりに

建設現場の映像から学習データを作成し、建機検出 AI とカテゴリ認識 AI に建設機械の外観と名称を学習させることで、建設機械を自動的に検出するシステムを開発した。このシステムは、適切な条件下であれば映像中の建設機械を±14%程度の誤差で検出できることがわかり、現場の出来高管理に適用可能である。その結果として、職員が現場に長時間留まる必要がなくなり効率的に業務を進めることができる。

このような有用性が明らかとなった一方で、映像と AI を用いたシステムを運用する場合の留意点もいくつか判明した。大きく分類すると、映像の撮影方法、学習データの収集・作成、AI の検出精度の 3 つに分類され、これらの課題を解決するに

はシステムの改良だけでは限界があり、現場との密接な連携が必要となる。具体的には、現場条件に合致するカメラの選定や実際に使用されている建設機械をもとにした学習データの作成がある。しかしながら、特定の現場に特化したシステムとした場合は汎用性が新たな課題となる。このシステムを展開し、現場の効率的な管理を実現させるためには様々な現場状況に対応する必要がある、今後はシステムに汎用性を持たせる技術開発が求められる。そのために、汎用的な学習データの作成方法やより少ない学習データで検出精度を向上させる手法についてさらに検討していきたい。

参考文献

- 1) 佐藤 諒, 木付拓磨, 松本江基, 早川健太郎, 花坂弘之, 朝倉健介, 伊藤祐介, 蒔苗耕司: 4K 定点カメラ映像による工事進捗管理システムの開発および試行結果(その1), 土木学会第75回年次学術講演会, VI-990, 2020.
- 2) 木付拓磨, 佐藤 諒, 松本江基, 早川健太郎, 永井利幸, 増田裕正, 伊藤祐介, 蒔苗耕司: 4K 定点カメラ映像による工事進捗管理システムの開発および試行結果(その2), 土木学会第75回年次学術講演会, VI-991, 2020.
- 3) Chen Sun, Abhinav Shrivastava, Saurabh Singh, Abhinav Gupta : Revisiting Unreasonable Effectiveness of Data in Deep Learning Era, In ICCV, 2017.
- 4) 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, AI・データサイエンスシンポジウム(投稿中), 2020.