

深層学習モデルを用いた画像による作業員検出評価

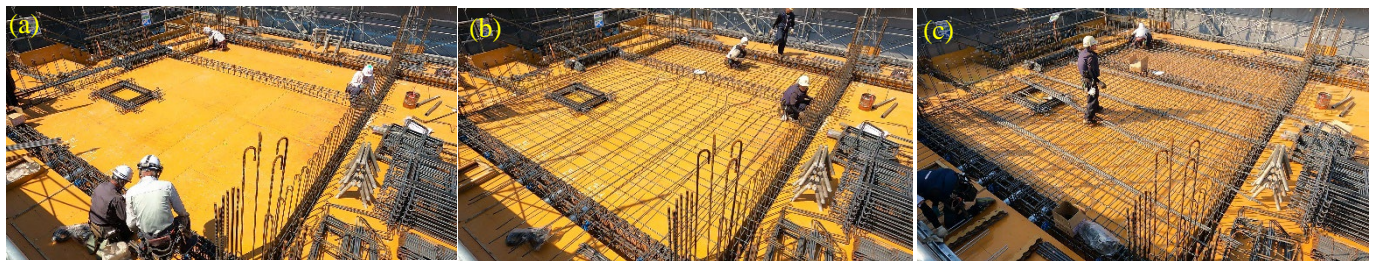
鹿島建設(株) 正会員 ○北原靖之 大塩 真 片村立太 露木健一郎 後閑淳司 中村泰広

1. はじめに

建設現場では、時間外労働削減や生産性向上の観点から、安全管理・進捗管理に関する省力化が求められている。最近では監視カメラの導入が進み、現場監督者が事務所に在席したまま作業状況をモニタリングできるようになりつつあるが、さらなる作業監視の省力化技術として画像認識技術の適用を検討している。近年急速に進歩している深層学習型物体検出を用いて、カメラ映像から作業員を検出し、作業中であるか否か自動で判断できれば省力化につながると考える。本稿では、物体検出モデルである YOLOv3, YOLOv4, YOLOv5x の 3 モデルを用いて実際の建設作業画像に対して人体検出を試み、性能を比較評価した結果を報告する。

2. 検証用画像と適用モデル

本検討では、躯体構築におけるスラブ配筋作業を対象とした。作業エリアに対してアクションカメラ (GoPRO9, 4K) を俯瞰する角度で作業足場に設置し、特定のスラブ面 (6m×8m) における一連の作業について開始から終了まで撮影した。取得した約 3 時間程度の作業動画からランダムにフレームを抽出し 121 枚の静止画像を作成した。画像例を写真-1 に示す。これらの画像に対して、物体検出モデルによる作業員検出を実行した。本検討では、物体検出における代表的なモデルである YOLO のうち、比較的新しい 3 モデル YOLOv3¹⁾, YOLOv4²⁾, YOLOv5x³⁾ について比較を試みた。3 モデルとも Microsoft COCO データセットによる学習済のモデルを使用した。検出結果とは別途、目視で画像中における作業員を確認し正誤を比較した。



(a) 撮影開始直後

(b) 撮影開始 36 分後

(c) 撮影開始 124 分後

写真-1 検証に用いたスラブ配筋作業の画像例

3. 検証結果

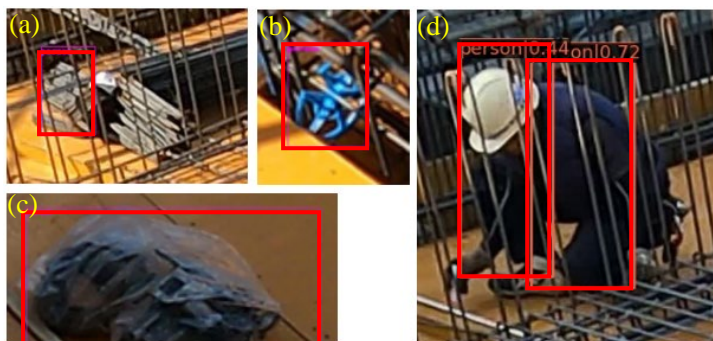
定量比較を行うため、各画像に対する検出結果に基づいてモデル毎に混合行列を作成し、適合率、再現率、F 値を算出した。結果を表-1 に示す。ここで、適合率はモデルによる検出結果のうち、正しく作業員が検出された割合を示す指標である。また、再現率は画像に含まれる全ての作業員のうち、モデルにより正しく検出された割合を示す指標である。F 値は適合率と再現率の調和平均値であり、両指標を総合的に判断する指標である。F 値で比較すると YOLOv5x が最も高い値となったが、適合率で比較すると YOLOv3 が最も高く、再現率で比較すると YOLOv4 が最も高い値となった。これらの結果から、YOLOv5x が適合率と再現率を総合的に判断して最も優れた検出性能を有するといえる。また、いずれのモデルも F 値で 80%~90% 程度の値となり、一定程度の検出性能を期待できるものの、安全管理等高い信頼性が要求される場合には、検出性能に課題があることが判明した。ただし、各指標値の差異がモデル間でそれほど大きくないため、検出対象となる作業内容や作業環境および検出閾値の最適化により、モデルの優劣の結果が逆転する可能性があると考えられる。安全管理等の作業員検出の確実性を重視する場合は、再現率が高い YOLOv4 を選択すべき場合もあると考える。

キーワード 画像認識, 深層学習, 物体検出, 人体検出

連絡先 〒182-0036 東京都調布市飛田給 2-19-1 鹿島建設(株) 技術研究所 TEL 090-485-1111

表-1 物体検出モデルによる作業員検出性能の比較一覧

モデル	適合率(%)	再現率(%)	F 値(%)
YOLOv3	99.1	74	85
YOLOv4	86	87	87
YOLOv5x	98.7	84	91



(a) 配筋前鉄筋の一部 (b) 梁用固定済スペーサー
(c) 資材用袋 (d) 同一人物を複数人物として検出

写真-2 偽陽性の例



(a) カメラに背を向けたしゃがみ姿勢
(b) 複数人物を単一人物として検出

写真-3 偽陰性の例

次に写真-2 に偽陽性となった例を、写真-3 に偽陰性となった例を示す。偽陽性例は、人体でない物体が人体として検出された例であり、偽陰性例は、人体が正しく検出されなかった例である。偽陽性に関しては現場にある資材が検出されており、人体と異なる特徴のある物体も検出される場合があることを確認した。また、同一人物を部位毎に複数人として検出する場合があることを確認した。偽陰性に関してはカメラに背を向けたしゃがみ姿勢において未検出となる場合があること、複数の人物が錯綜する状態で複数人物を単一人物として検出される場合があることを確認した。この未検出要因としては、学習に使用したデータセットにおいて類似の姿勢や状態が十分に含まれていない可能性があると考えられる。検出性能向上のためには、現場作業員画像を用いた追加学習が必要となると考える。

4. まとめ

現場監視の省力化を目的として、学習済み物体検出モデルを用いた建設現場における作業員検出を実施し、YOLOv3, YOLOv4, YOLOv5x の検出結果を比較評価した。今回比較した条件においては、適合率と再現率を総合的に評価して YOLOv5x が最も優れた検出結果となった。また、いずれのモデルにおいても一定水準での検出性能が得られることを確認したが、安全管理等高い信頼性が要求される場合には検出性能に課題があることを確認した。今後は、検出されなかった姿勢や状態画像を用いた追加学習による性能向上や、異なるカメラ設置条件および他工種作業における性能評価などを進めていく予定である。

参考文献および Web サイト

- 1) "YOLOv3: An Incremental Improvement", Joseph Redmon, Ali Farhadi, (CVPR, 2018), <https://arxiv.org/abs/1804.02767v1>.
- 2) "YOLOv4: Optimal Speed and Accuracy of Object Detection", Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao (CVPR, 2020), <https://arxiv.org/abs/2004.10934v1>.
- 3) YOLOv5: Glenn Jocher, <https://github.com/ultralytics/yolov5> (released 2020).