

## (22) 建設現場における人物検出のための 深層学習用学習データの 自動生成に関する研究

松尾 龍平<sup>1</sup>・姜 文淵<sup>2</sup>・山本 雄平<sup>3</sup>・中村 健二<sup>4</sup>・  
田中 ちひろ<sup>5</sup>・田中 成典<sup>6</sup>・鳴尾 丈司<sup>7</sup>

<sup>1</sup>学生会員 関西大学大学院 総合情報学研究科 (〒569-1095 大阪府高槻市霊仙寺町 2-1-1)  
E-mail: k133395@kansai-u.ac.jp

<sup>2</sup>正会員 関西大学特別任命准教授 先端科学技術推進機構 (〒564-8680 大阪府吹田市山手町 3-3-35)  
E-mail: b.kyo@kansai-u.ac.jp

<sup>3</sup>正会員 関西大学助教 環境都市工学部 (〒564-8680 大阪府吹田市山手町 3-3-35)  
E-mail: y.yamamo@kansai-u.ac.jp

<sup>4</sup>正会員 大阪経済大学教授 情報社会学部 (〒533-8533 大阪府大阪市東淀川区大隅 2-2-8)  
E-mail: k-nakamu@osaka-ue.ac.jp

<sup>5</sup>非会員 関西大学特別任命助教 先端科学技術推進機構 (〒564-8680 大阪府吹田市山手町 3-3-35)  
E-mail: c\_tanaka@kansai-u.ac.jp

<sup>6</sup>正会員 関西大学教授 総合情報学部 (〒569-1095 大阪府高槻市霊仙寺町 2-1-1)  
E-mail: tanaka@res.kutc.kansai-u.ac.jp

<sup>7</sup>非会員 関西大学特別任命教授 先端科学技術推進機構 (〒564-8680 大阪府吹田市山手町 3-3-35)  
E-mail: tnaruo@kansai-u.ac.jp

建設現場における労働災害をゼロにするには、効果的な安全管理の対策の徹底が肝要である。近年、動画像から人物や物体を自動で識別する物体検出手法が様々な分野で技術発展を遂げている。作業員や建設機械が混在して稼働し、危険な領域が常時変化する建設現場において、動画像から物体検出手法を用いて人物を検出をすることで安全管理に貢献できる。しかし、作業員を正確に検出するためには、建設現場に特化した検出モデルへ更新する必要があるが、学習データをマニュアルで作成するには多大な労力を要する。そこで、本研究では、建設現場における人物検出のための深層学習用学習データの自動生成手法を提案する。遠距離から撮影した動画像を建設現場全体を撮影した映像と見立てて施行した結果、適用可能性のある示唆を得ることができた。

**Key Words:** construction site, people identification, deep learning, automatical generation on training data

### 1. はじめに

近年、IoT の急速な普及に伴い、様々な機器によって人物の動きを解析し、産業や事業に活用する取り組みが普及している。この潮流を加速させるため、日本政府は科学技術基本計画<sup>1)</sup>にて、IoT、ロボット、人工知能やビッグデータなどの先端技術を駆使した経済発展と社会課題の解決を目指した Society5.0 を提唱している。

これを受け、カメラで撮影された映像中の人物や物体

を人工知能により自動で識別する物体検出手法を用いた社会発展を促進する取り組みが行われている。具体例として、監視カメラにより撮影された映像から人物を自動検出する例<sup>2)</sup>や、自動運転時に周囲の状況の検知に用いる例<sup>3)</sup>など、物体検出手法は多岐にわたって利用されている。この技術動向により、建設分野において高精度に人物を検出することができれば、生産性や安全管理の発展に大きく寄与することが期待できる。

建設業労働災害防止協会によると、建設業による死傷

者数は年々減少傾向にあるものの、2020年には約15,000人が死傷者となっている。その原因として、建設機械（以下、建機）の転倒、倒壊、墜落、資材の飛来・落下などの事故が挙げられる。労働災害を防止するには、様々な技術を駆使した効果的な安全管理の追求と徹底が求められる。こうした背景の下、著者らは、多くの作業員や建機が混在して稼働し、危険な領域が常時変化する建設現場において、撮影した動画画像から物体検出手法を用いて人物を検出することで安全管理に貢献できると考えた。具体的には、人物を検出することで、墜落や転倒の可能性がある危険箇所への侵入、建機との異常接近などを検知することが可能となる。

建設現場の映像に物体検出手法を適用した場合、物体同士が重なり合うオクルージョンが発生することや、撮影角度、アングルにより作業員が建機で隠れるなど、正しく人物を検出できない課題がある。そのため、学習データを作成し、検出モデルを建設現場に特化したモデルへ更新する必要があるが、学習データをマニュアルで作成するには、多大な労力を要する。そこで、本研究では、建設現場において、作業員を検出するための深層学習用の学習データを自動生成する手法を提案する。これにより、建設現場の作業員の位置や動線を容易に解析でき、低コストで、現場全体の安全性を確保できると考えられる。

本研究では、まず、2章で学習データ生成技術について調査し、3章で本研究の目的を明示し、提案手法について詳述する。そして、4章にて実験を行い、最後に、5章にて本研究の総括を行う。

## 2. 既存研究の調査とその考察

深層学習用の学習データ生成技術として、類似度が高い画像を収集する研究<sup>5)</sup>や、身体パーツを検出し学習データとする研究<sup>6)</sup>がある。

既存研究<sup>5)</sup>の技術では、準備されたサンプル画像との類似度が高い画像を学習データとして抽出する。しかし、建設現場にこの方法を適用した場合、人物の特徴量が少なく、学習データとして取得できない。

また、既存研究<sup>6)</sup>では、DPM (Deformable Part Model) 技術を用いて身体パーツの検出を行い、人物を検出している。この手法を適用して学習データとすることが考えられるが、建設現場での動画画像に対して適用した場合、人物が小さく撮影され、身体パーツの取得が困難なことから、検出結果に多くの誤検出や検出漏れがあり、そのまま学習データとして用いると、検出精度が低下する。

これらの課題を鑑み、本研究では、建設現場の映像から既存の検出モデルにより検出した結果を手作業に頼ら

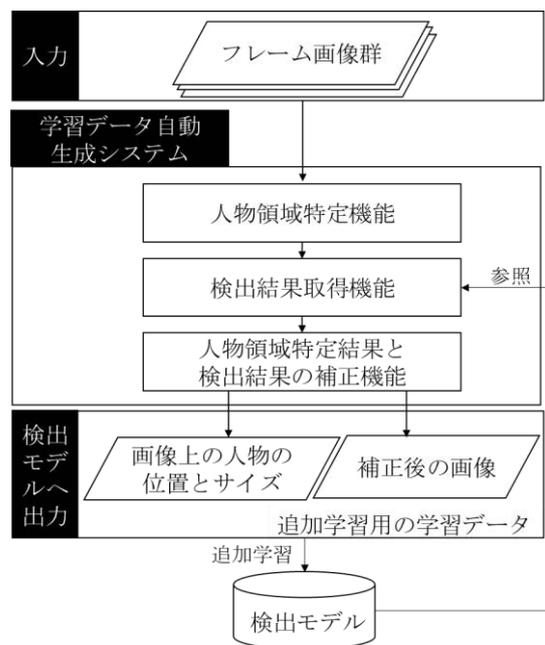


図-1 提案手法の処理フロー

ず自動で精査し、現場に特化した新たな学習データとすることで、この課題解決を目指す。

## 3. 建設現場における学習データの自動生成手法

### (1) 研究目的と内容

本研究では、人物の検出が困難な建設現場の映像に対して、深層学習用の学習データ自動生成手法を提案する。この提案手法により生成した学習データを用いて追加学習を行うことで、検出モデルの精度の向上を目指す。これにより、基礎工事や、躯体工事、外装・内装工事など、異なる建設現場にてそれぞれの映像に適切な検出モデルの構築を実現する。

### (2) 研究目的と内容

提案手法の処理フローを図-1に示す。提案手法では、人物領域特定機能、検出結果取得機能、人物領域特定結果と検出結果の補正機能により構成される。入力データは、建設現場全体を撮影した映像から切り出したフレーム画像群とする。出力は、追加学習用の学習データとなる補正された人物の位置とサイズ及び補正後の画像群とする。

### (3) 人物領域特定機能

本機能では、背景差分法を用いて、作業員が存在する可能性がある領域（以下、人物領域）を特定する。しかし、建設現場では作業員は常に移動しており、日照条件も刻々と変化することから、良好な背景画像を取得する

ことが困難であると予想される。そこで、本研究では、OpenCV<sup>7)</sup>にライブラリ化されている混合正規分布に基づいた背景とそれ以外を分離する手法<sup>8)</sup>を用いて自動的に背景画像を生成する。

#### (4) 検出結果取得機能

本機能では、MS-COCO<sup>9)</sup>データセットにて学習済みのモデル（以下、既存モデル）を参照して人物を検出し、その結果を取得する。これにより、映像の人物領域の位置とサイズを取得できる。既存モデルでは、80クラスの物体を検出できるが、本実験では、人物だけを用いる。

検出においては、深層学習の一種であり高度な画像解析能力を持つ畳み込みニューラルネットワーク（CNN：Convolutional Neural Network）を有した物体検出・認識アルゴリズムであるYOLOv3<sup>10)</sup>を使用する。

#### (5) 人物領域特定結果と検出結果の補正機能

本機能では、人物領域特定結果と検出結果を重畳することで、誤検出や人物領域特定時に発生するノイズを除去する。まず、人物領域と検出結果を重畳する。次に、検出結果の矩形内を占める人物領域の割合を算出する。最後に、矩形内を占める領域の割合が一定以上の場合、検出結果は正しいと判断し、その位置とサイズを学習データとして抽出する。検出結果の矩形内を占める人物領域の割合が一定以下の場合、この検出結果は誤検出であると判断し、この検出結果を黒く塗りつぶす処理を行う。これにより、誤りの学習データを除去することで、質の向上を目指す。

## 4. 実証実験

### (1) 実験内容

本実験では、建設現場で撮影された映像を想定し、映像内の人物領域が小さく撮影された映像を対象に検証を行う。実験に用いる映像データは、関西大学高槻キャンパスにて、建設現場を想定した行動を被験者10名が行い、それを撮影したものを用いる。作業員を再現するために、被験者には、ヘルメットを着用してもらった。また、撮影にはスポーツコーチングカメラ（JVCケンウッド社、2018）を使用し、Full-HD、60fpsで撮影を行った。この映像で撮影された人物のサイズは平均35pix程度であった。使用した映像のイメージを図-2に示す。

実験では、既存モデルと、提案手法により生成した学習データで追加学習したモデルの双方の精度を比較することで、提案手法の有用性を検証する。検証時には、図-2の映像にて学習時に使用していないフレームから取得した100フレームを評価データとして用いる。精度を評価



図-2 本研究に使用した映像のイメージ



図-3 生成した学習データの例

するため、各モデルに適用して、人物検出を行う。検出結果から検出された人数、誤検出された人数をカウントし、再現率、適合率、F値を算出し評価を行う。

### (2) 学習データ

学習データは、図-2の映像から取得した2,600枚のフレーム画像群を提案手法の学習データ自動生成システムに入力し、得られた画像群及び人物の位置データとした学習データの一例を図-3に示す。

提案手法によって誤検出結果を黒く塗りつぶす処理が行われていることが確認できる。

### (3) 実験結果

各検出器の検出結果の100枚の平均を表-1に示す。まず、既存モデルでは、適合率の平均が0.70、再現率の平均が0.35、F値の平均が0.45となり全体的に低い値となった。これは、画像内の人物が小さく撮影されていることや、オクルージョンが頻繁に発生していることが原因と考えられる。結果を詳細に確認すると、図-4に示すように、A箇所では、検出に失敗している例や、B箇所では、誤った位置での検出結果も見られた。これは、人物のサイズが小さく、既存モデルではこのような特徴を学習していないためであると考えられる。これは、近似する特徴を学習することで対応が可能である。

次に、追加学習を行った検出モデルでは、適合率の平均が0.81、再現率の平均が0.48、F値の平均が0.59となった。既存モデルと比較すると、精度が向上したことがわかった。これは、この映像内の人物の特徴をより正確に捉えることができ、精度よく検出が可能になったと考

表-1 検出結果

	映像内の人数	既存モデル					追加学習した検出モデル				
		検出された人数	誤検出された人数	適合率	再現率	F値	検出された人数	誤検出された人数	適合率	再現率	F値
100枚平均	9.13	3.21	0.78	0.70	0.35	0.45	4.36	0.89	0.81	0.48	0.59

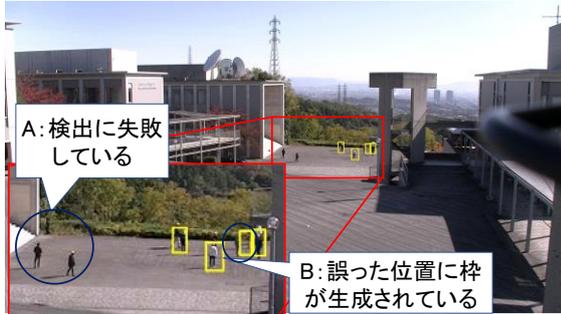


図-4 既存モデルでの検出結果例

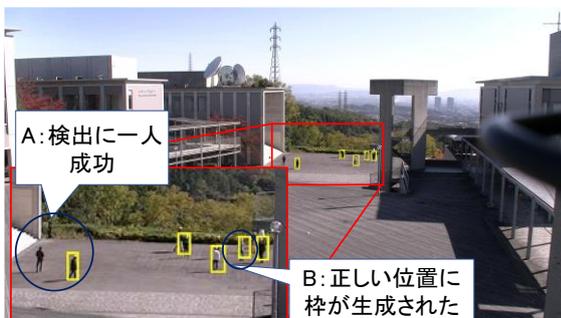


図-5 追加学習した検出モデルでの検出結果例

えられる。図-4同様のフレームでの結果を図-5に示す。図-5を確認すると、既存モデルでは、検出ができなかったA箇所にて、人物の検出に成功していることがわかる。また、B箇所では、誤った位置での結果が補正されている。このように提案手法を用いて追加学習を行った結果、検出精度が向上したことから、提案手法は有用であることがわかった。一方、検出できていない人物も確認できるが、提案手法を用いて検出結果を学習データとし検出器を繰り返し更新することで、精度向上できると考えられる。

## 5. おわりに

本研究では、建設現場における作業者を検出するための深層学習用の学習データの自動生成手法を考案した。建設現場を想定した映像を用いて追加学習及び検出を試行した結果、既存モデルに比べて検出精度が向上した。これにより、汎用性のある既存モデルから簡便に映像ごとに専用のモデルの生成ができる。

今後は、実現場を想定した環境下において、精度検証を重ねることにより、提案手法の洗練を図る。また、人物のみならず建機を検出するための学習データの自動生成手法を提案することで、より現場の安全性に寄与していく予定である。

## 参考文献

- 1) 内閣府：科学技術基本計画、<<https://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf>>、（入手 2021.6.14）。
- 2) 矢野光太郎，河合智明：監視カメラにおける映像認識の動向，日本画像学会誌，Vol.55，No.3，pp.341-347，2016。
- 3) 内閣官房 IT 総合戦略室：自動運転・MaaS を巡る最近の動向、<[https://www.kantei.go.jp/jp/singi/it2/dourokoutsu\\_wg/dai12/sankou3.pdf](https://www.kantei.go.jp/jp/singi/it2/dourokoutsu_wg/dai12/sankou3.pdf)>、（入手 2021.6.14）。
- 4) 建設業労働災害防止協会：建設業における労働災害発生状況、<[https://www.kensaibou.or.jp/safe\\_tech/statistics/occupational\\_accidents.html](https://www.kensaibou.or.jp/safe_tech/statistics/occupational_accidents.html)>、（入手 2021.6.14）。
- 5) Das, N., Chaba, S., Wu, R., Gandhi, S., Chau, D. and Chu, X.: GOGGLES: Automatic Training Data Generation with Affinity Coding, *International Conference on Management of Data*, pp.1717-1732, 2020.
- 6) Lu, C., Hsu, C., Kang, L., Lin, C., Weng, M. and Liao, H.: Identification and Tracking of Players in Sport Videos, *Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service*, pp.113-116, 2013.
- 7) OpenCV team : OpenCV 3.0, <<https://opencv.org/opencv-3-0/>>、（入手 2021.6.14）。
- 8) Zivkovic, Z. and Heijden, F.: Efficient Adaptive Density Estimation Per Image Pixel for the Task of Background Subtraction, *Pattern Recognition Letters*, Vol.27, No.7, pp.773-780, 2006.
- 9) Lin, T., Maire, M., Belongie, S., Hays, J. and Perona, P.: Microsoft COCO: Common Objects in Context, *European Conference on Computer Vision*, pp. 740-755, 2014.
- 10) Redmon, J., Divvala, S., Girshick, R. and Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection, *Computer Vision and Pattern Recognition*, Vol.29, No.2, pp.779-788, 2016.