

# 建設現場の人物認識のための YOLO の特性分析に関する研究

井上晴可<sup>†</sup> 梅原喜政<sup>‡</sup> 田中成典<sup>††</sup> 今井龍一<sup>†‡</sup> 神谷大介<sup>††</sup> 中畑光貴<sup>‡‡</sup> 丹野豪<sup>††</sup>

大阪経済大学情報社会学部<sup>†</sup> 関西大学先端科学技術推進機構<sup>‡</sup> 関西大学総合情報学部<sup>††</sup>

法政大学デザイン工学部<sup>†‡</sup> 琉球大学工学部<sup>††</sup> 関西大学大学院総合情報学研究科<sup>‡‡</sup>

## 1. はじめに

近年、建設現場において甚大な事故が多数発生[1]しており、情報技術を活用した安全管理の方法が求められている。例えば、作業員の行動を把握して、潜在的な危険を顕在化するヒヤリハットの検知技術により、事故の発生を未然に防止できると考えられる。その根幹を担う技術として、人物の位置を特定する方法が重要となる。既存研究[2]では、カメラの動画から深層学習により、識別した人物の位置を特定する手法が提案されている。しかし、深層学習は出力結果に至った根拠を説明できず、また失敗した原因を分析できないことから、人物認識に適した撮影環境や計測条件が定まらない。そこで、本研究では、既存研究の物体検出手法である YOLO (You Only Look Once) [3]がどのような撮影環境下で人物を認識できるか（以下、人物認識可否）や人物認識時に人のどの部位に着目しているか（以下、着目箇所）を明らかにする。これにより、建設現場における人物認識に適したカメラの撮影要件を提案する。

## 2. YOLO の特性分析手法

本研究の分析フローを図 1 に示す。取り扱うデータは、建設現場を想定して撮影した作業員の動画として、分析目標は建設現場における人物認識可否と着目箇所、そして、撮影要件とする。

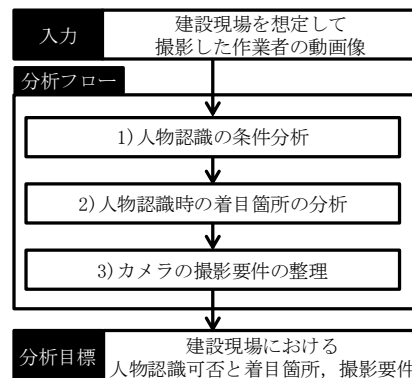


図1 分析フロー

表 1 YOLO の人物認識精度に関わる要因

| 影響要因  | 条件値             |
|-------|-----------------|
| カメラ角度 | 0度, 30度, 45度    |
| 撮影距離  | 5m, 10m, 15m    |
| 背景    | あり, なし          |
| 明度    | 明るい, 暗い         |
| 向き    | 前後左右            |
| 姿勢    | 直立, 前屈, 屈伸      |
| 部位の消去 | なし, 頭, 胴体, 下半身  |
| 人物の色  | 変更なし, 単色, 動物の模様 |

### 2. 1 人物認識の条件分析

本項目では、まず、YOLO の人物認識精度に関わると考えられる要因を整理（表 1）する。次に、表 1 に基づいて撮影した動画に対して YOLO を適用する。最後に、その認識結果を確認して、人物認識可否を分析する。

### 2. 2 人物認識時の着目箇所の分析

本項目では、2.1 節の人物認識に成功した画像に対して、着目箇所をヒートマップで可視化するために Grad-CAM[4]を適用する。これにより、影響要因ごとの着目箇所を分析する。

### 2. 3 カメラの撮影要件の整理

本項目では、2.1～2.2 節の結果から対応関係を整理する。そして、その結果に基づいて建設現場での人物認識に適したカメラの撮影要件を明らかにする。

Research Concerning Characteristic Analysis of YOLO for Person Identification at Construction Sites

<sup>†</sup> Haruka Inoue

Faculty of Information Technology and Social Sciences, Osaka University of Economics

<sup>‡</sup> Yoshimasa Umehara

Organization for Research and Development of Innovative Science and Technology, Kansai University

<sup>††</sup> Shigenori Tanaka and Tsuyoshi Tanno

Faculty of Informatics, Kansai University

<sup>†‡</sup> Ryuichi Imai

Faculty of Engineering and Design, Hosei University

<sup>‡‡</sup> Daisuke Kamiya

Faculty of Engineering, University of the Ryukyus

<sup>††‡</sup> Koki Nakahata

Graduate School of Informatics, Kansai University

### 3. 実証実験

実証実験では、建設現場を想定した画像を対象に、本研究の分析手法を適用することでYOLOを分析する。

#### 3.1 実験内容

本実験では、ヘルメットと安全ベストを着用した人物を表1に示す条件に基づいて関西大学高槻キャンパス内にて撮影する。次に、条件ごとに3画像ずつ選定し、背景、明度、部位の消去と人物の色は、図2を元画像として条件に合うように加工する。そして、分析フローに従って、人物認識可否や着目箇所を明らかにすることで、最終的に撮影要件をまとめる。なお、カメラはソニー社の4Kカメラ(FDR-AX45)を用いた。

#### 3.2 結果と考察

実験結果を表2に示す。認識可否を確認すると、カメラ角度を水平方向から撮影した場合や、撮影距離を近距離で撮影した場合は人物を認識できるが、カメラ角度が大きくなる場合や、撮影距離が遠くなる場合に人物を認識できなかった。これは、YOLOの学習データ(COCO[5])には、正対した人物画像や人物領域が大きい画像が多く含まれているためと考えられる。

また、表2の着目箇所を確認すると、胸部、腹部や足に着目して人物を認識している傾向があることがわかった。着目箇所の可視化結果を確認すると、図3に示すとおり、胸部、腹部や足を局所的に着目しており、これらの箇所に人物としての特徴が多く表れていることを確認した。

以上のことから、YOLOを用いた人物認識時の撮影要件は、カメラを人物に対して正対する角度で、人物が大きく写るように拡大して撮影し、胸部、腹部や足が隠れないように設置位置を工夫する必要があることがわかった。

#### 4. おわりに

本研究では、YOLOを対象として、人物認識可否や着目箇所を整理することで、建設現場での人物認識に適したカメラの撮影要件を明らかにした。今後は、実際の建設現場においても本検証の手法を適用する予定である。

#### 参考文献

- [1] 建設労働災害防止協会：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2021.1.8)。
- [2] 今井龍一、神谷大介、井上晴可、田中成典、坂本一磨、藤井琢哉、菊地英一、伊藤誠：深層学習を用いた建設現場における人物認識に関する研究、土木学会論文集 F3 (土木情報学)、土木学会、Vol.75, No.2, pp. I\_57-I\_68, 2019。
- [3] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. : You Only Look Once: Unified, Real-Time Object Detection,

表2 認識結果 (一部抜粋)

| 影響要因  | 条件値 | 認識可否 | 着目箇所  |
|-------|-----|------|-------|
| カメラ角度 | 0度  | ○    | 足     |
|       | 30度 | ×    | -     |
|       | 45度 | ×    | -     |
| 撮影距離  | 5m  | ○    | 腹部    |
|       | 10m | ×    | -     |
|       | 15m | ×    | -     |
| 背景    | あり  | ○    | 足     |
|       | なし  | ○    | 腹部    |
| 明度    | 明るい | ○    | 足     |
|       | 暗い  | ○    | 足     |
| 部位の消去 | なし  | ○    | 足     |
|       | 頭   | ○    | 腹部, 足 |
|       | 胴体  | ○    | 頭部, 足 |
|       | 下半身 | ○    | 胸部, 足 |



図2 元画像

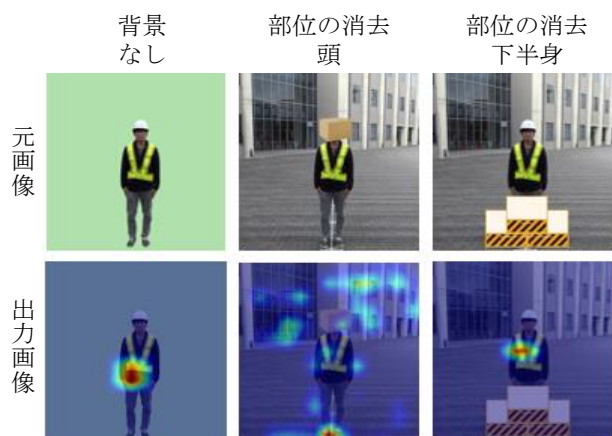


図3 着目箇所の可視化結果

Computer Vision and Pattern Recognition, IEEE, Vol.29, No.2, pp.779-788, 2016.

- [4] Selvaraju, R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. : Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, <<https://arxiv.org/abs/1610.02391>>, (入手2021.1.8)。
- [5] COCO : Common Objects in Context, <<https://cocodataset.org/#home>>, (入手2021.1.8)。