

周年栽培トマトの果実検出システムの検証

Crop Yield Estimation for Hydroponic Tomatoes using regional CNNs

磯山 侑里^{*1} 江村 文代^{*1} 佐藤 裕久^{*1} 篠崎 隆志^{*2*3}
Yuri Isoyama Fumiyo Emura Hirohisa Satoh Takashi Shinozaki

^{*1} 協和株式会社
Kyowa Co. Ltd.,

^{*2} 国立研究開発法人情報通信研究機構脳情報通信融合研究センター
Japanese Society for Artificial IntelligenceCiNet, National Institute of Information and Communications Technology

^{*3} 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

Tomatoes are important plants in greenhouse cultivation from the rise in their agricultural productivity by spread of environment management system, and its crop yield estimation is crucial for efficient cultivation and controlled distribution. We develop a crop yield estimation system for hydroponic tomatoes using regional convolutional neural networks and verify the detection score and estimation accuracy.

1. はじめに

農業は近年、計画生産、計画出荷の需要が高まっている。特に施設栽培のトマトでは収穫量の増減による機会損失や収穫物の廃棄が課題になっており、中長期的な収穫量の予測は喫緊の課題となっている。一週間先の収穫量を予測するためには果房についている果実数、およびその状態を把握することが好ましいが、現状では栽培管理者が目視で判断しており、定量的な評価が行われていないのが実状である。

近年の人工知能技術は、特に画像処理分野での進歩が著しく、畳み込みニューラルネットワーク(Convolutional Neural Network, 以下 CNN) による画像分類を発端に、画像検出に応用した Regional CNN(以下 RCNN)と呼ばれる技術が急速に発展してきている。RCNN を農業に応用した先行研究としては、リンゴの果実検出を行った報告 [小林 18] がなされているが、より高い空間密度で結実するトマトに関しては十分な研究が行われていない。そこで本研究では栽培中のトマトの画像から翌週の収穫量を予測することを目的として、様々な分野で有効性が確認されつつある RCNN である Faster RCNN (以下 F-RCNN) [Ren 15] および Single Shot Detector (以下 SSD) [Liu 16] を用いたトマトの果実検出システムを構築し、その検出精度と収穫量予測における精度を検証した。

2. 手法

2.1 データセット

2018 年 10 月から 2019 年 2 月の間に周年栽培のミニトマトの収穫時に果実が含まれる果房をデジタルカメラで撮影した。撮影した画像(4608×3456 ピクセル)から、トマトが結実している高さで 1600×1200 ピクセルの画像を 3 枚切り出し、800×600 ピクセルに縮小した。切り出された画像に映っている果実を、専用に作成されたタグ付けソフトウェアを用いて、収穫作業者が赤(収穫できる状態)、オレンジ(成熟途中)、緑(未成熟)の 3 種類にタグ付けした(図 1)。画像の総数は 107 枚であり、73 枚を学習用、25 枚を検証用、9 枚を評価用に分割して用いた。

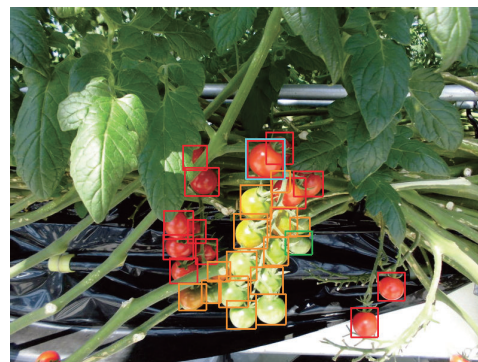


図 1. タグ付けされた撮影画像

2.2 RCNN

RCNN の実装には Preferred Networks 社製の Chainer [Tokui 15] および ChainerCV [Niitani 17] を用いた。学習および評価の各過程は NVIDIA 社製 GeForce 1070Ti を搭載したワークステーションによって行った。

F-RCNN では 800×600 ピクセルの画像を入力し、ミニバッチサイズを 1 として、3000 試行の学習を行った。SSD では 800×600 ピクセルの画像から中央の 512×512 ピクセルを切り出した後、ミニバッチサイズを 8 として、1500 試行の学習を行った。

学習は PASCAL VOC で学習済みのモデルからの転移学習によって行った。重みの更新には通常の Stochastic Gradient Descent (以下 SGD)を用い、学習係数は初期値を 0.01 とし、全試行の 1/3 ごとに学習係数を 1/10 にしたものをを用いた。画像検出の精度は検証用データに対する mean Average Precision (以下 mAP)によって確認し、収穫量予測には評価用データに対して 10 分割交差検証を行い、評価用の 9 枚の画像毎、およびに画像全体を通しての検出数を確認した。

3. 実験及び結果

図 2 に学習過程における損失、およびに検証用画像に対する mAP の一例を示す。10 試行の学習に対する mAP は F-RCNN が 52.9±2.7%, SSD が 63.3±1.5%であった。1 試行当たりの学習に要した時間は FRCNN が 1190±6 秒、SSD が 1216±8 秒であった。

連絡先: 磯山侑里, 〒569-1136 大阪府高槻市郡家新町 85-1

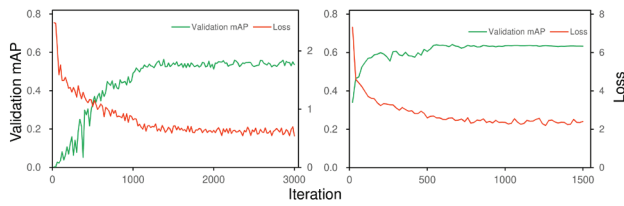


図 2. 学習過程の損失, およびに精度の変化

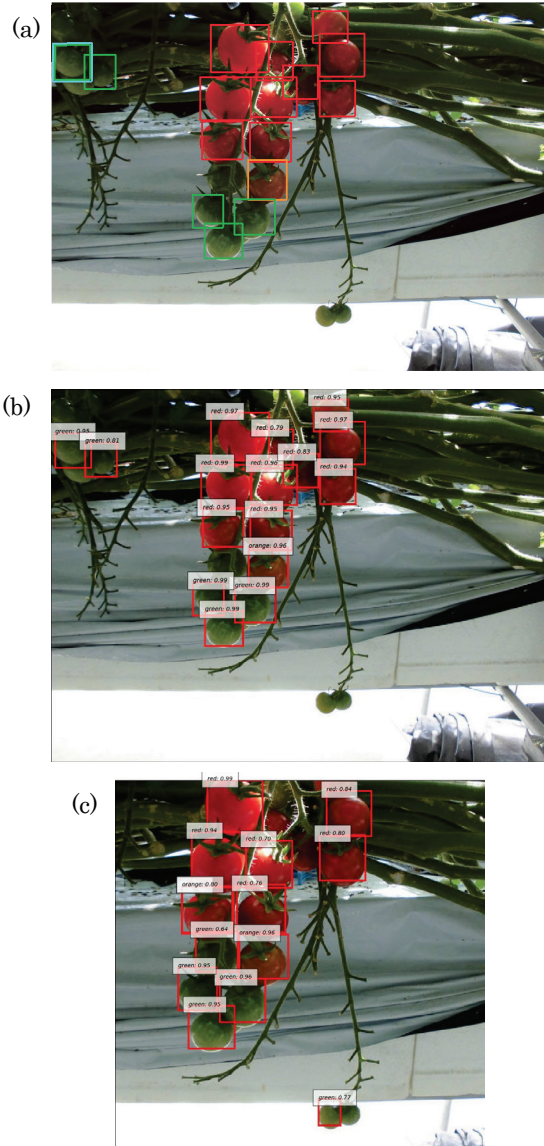


図 3. 学習済みモデルによる画像検出の結果. (a) 正解データ, (b) F-RCNN, (c) SSD

学習が完了したそれぞれのモデルを用いて, 評価用画像のトマト果実検出を行った結果を図 3 に示す. 図 3(a)が正解データ, 図 3(b)が F-RCNN での検出結果, 図 3(c)が SSD での検出結果となっている. 図の様に赤, オレンジ, 緑の果実を十分な精度で検出できることが確認された.

果実検出の結果から, それぞれの手法における赤, オレンジ, 緑の果実の収穫量予測を, 画像毎およびに評価画像全体を通しての検出数によって行った. 表 1 に画像毎の収穫量予測の一例として, 代表的な試行における予測結果の画像間平均の結

表 1. 画像毎の収穫量予測の 1 例(%) (n=9)

	赤	オレンジ	緑
F-RCNN	94.7±24.2	94.9±36.3	88.9±33.3
SSD	109.2±41.5	78.3±55.5	79.5±40.0

表 2. 画像全体を通しての収穫量予測(%) (n=10)

	赤	オレンジ	緑
F-RCNN	89.5±10.8	97.5±7.7	92.3±14.1
SSD	81.6±13.8	85.5±9.9	80.0±21.1

果を示す. SSD では赤とオレンジの判断間違いが F-RCNN より高い傾向にあり, 特にオレンジの果実を赤と判断することが多いため, 赤の果実の予測値が 100%を越える結果となった. 一方で F-RCNN では正解率の平均値自体は高い値になっているものの, 葉に隠れた場合の検出率が低くなる傾向があった. 表 2 に全画像を通しての収穫量予測の試行間平均の結果を示す. 1 枚ごとの評価に比べて, 特に検出数の標準偏差が低減され, 安定した精度での予測が可能であることが確認された.

4. 考察

本研究では独自に構成したトマト画像のデータセットを用いて, F-RCNN と SSD による周年栽培トマトの果実検出システムを構築, 検出精度と収穫量予測の精度についての検証を行った. その結果, SSD に比べて F-RCNN のほうがより安定した収穫量の予測が可能であった. さらに 1 枚毎の画像を用いた検出システムより複数の画像から検出した果実の総数で評価の方が検出数の精度が高かった. また, その標準偏差の小ささから農場全体の収穫量予測としては, 予測値に一定の係数を乗じることによって一定の精度の予測が可能であることが示唆された.

RCNN は現在も発展中の技術であり, 本研究で用いたデータセットや評価手法をより新しい RCNN に適用することによって更なる精度の向上が期待される. より高い精度の熟度の判定を可能とするためにカラーチャートを用いた色補正の導入などと併せて, さらなる改良を行ってゆく予定である.

5. 参考文献

- [Girshick 2014] Girshick, R., Donahue, J., Darrell, T. and Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014)
- [小林 18] 小林一樹, 中村俊輝: 深層学習を用いた果実生育情報抽出 Web アプリケーション, 第 32 回人工知能学会 (2018)
- [Ren 15] Ren, S., He, K., Girshick, R. and Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks, *Neural Information Processing Systems (NIPS)* (2015)
- [Liu 16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed S., Fu, C. and Berg, A.C.: SSD: Single shot multibox detector, *European Conference on Computer Vision (ECCV)* (2016)
- [Tokui 15] Tokui, S., Oono, K., Hido, S. and Clayton, J.: Chainer: a Next Generation Open Source Framework for Deep Learning, *Proc. Of Workshop on Machine Learning Systems (LearningSys)* (2012)
- [Niitani 17] Niitani, Y., Ogawa, T., Saito, S. Saito, M.: ChainerCV: a Library for Deep Learning in Computer Vision, *ACM Multimedia* (2017)