

農園画像における深層学習を用いた隠れ果実領域の抽出 Training Data Augmentation for Hidden Fruit Image Segmentation by using Deep Learning

高井 亮磨^{*1}
Ryoma Takai

小林 一樹^{*2}
Kazuki Kobayashi

^{*1} 信州大学 工学部
Faculty of engineering, Shinshu University

^{*2} 信州大学 学術研究院
Academic Assembly, Shinshu University

This paper proposes a method to detect apple fruits areas from field monitoring images. The proposed method can detect the fruits areas even if they are hidden by leaves or other fruits. We use a deep neural network with a way to create a large amounts of artificial field images as training data. Since the developed system uses the fruit and leaf image components and can easily retrieve their exact areas in the creation process, the created data include annotations of hidden fruits. The experimental results showed that the proposed method showed promising results in terms of accuracy compared to that of the COCO pre-trained model.

1. はじめに

近年, IoT や ICT を用いて農作物の生育状態を抽出し, 生産性の向上を目指す研究が行われている. たとえば, 深層学習を用いた果樹の果実検出 [Bargoti 2017] に関する研究では, 深層学習を利用した物体検出手法である Faster R-CNN [Ren 2015] を用いて, リンゴやマンゴー, アーモンドの農園画像から果実を検出している. これらの研究では果実を囲む矩形領域を検出している. もし果実領域を検出できれば, 正確なサイズと形状とを詳細に把握でき, 農園管理に有用である. しかし, 多くの果樹は果実を覆い隠すように葉を茂らせているため, 先行研究では葉で隠れた果実サイズの検出までは実現されていない.

そこで本研究では, 定点カメラで観測した農園モニタリング画像から, 果実の一部が隠れていても領域推定が可能な手法を提案する. 提案手法では, 深層学習を用いた隠れた果実領域抽出のための効率的な訓練データ拡張システムを開発する.

2. 隠れ果実領域の深層学習

深層学習での物体検出には大量の訓練データが必要であり, COCO [Lin 2014] のような既存の大規模データセットを用いて学習を行うことが多い. しかし, これらのデータセットには, 隠れた果実領域を検出するためのデータが存在しない. 大量の訓練データを人手で生成する作業には多大な労力を要するとともに, 隠れた領域に関しては, 人間が正確に領域を特定できない問題がある. そこで, 本研究では実際の農園モニタリング画像をベースとして, 人工的に大量の訓練データを生成するアプローチをとる.

2.1 果実領域の検出

農園観測画像から果実の領域を抽出するために, 畳み込みニューラルネットワークによる物体検出手法である Mask R-CNN [He 2017] を用いる. Mask R-CNN は, 画像の中の任意の対象について, その形状に合わせて領域を抽出する手法である.

本研究においては隠れている果実領域を推定するために, 果実が隠れている観測画像に対して, 隠れていない状態の果実領域を学習させる.

切り出し画像



自動生成画像



図 1: 合成画像の生成方法

2.2 訓練データ自動生成

生成する訓練データは, 合成農園画像と, 果実領域を示すバイナリマスクの 2 つである. 図 1 に合成農園画像の生成方法を示す. まず, 実際の農園モニタリング画像から手動で果実画像, 葉画像, 背景画像の切り出しを行う. このとき, 切り出された画像は PNG 形式の RGBA 画像とし, 果実画像と葉画像の背景は透明化されている. 次に, 切り抜いた画像の輝度や角度, 拡大率をランダムに変更したものをランダムな位置に配置することで, 合成農園画像を生成する.

バイナリマスクとは, 白黒の 2 値で表現された対象物の領域を表すデータである. Mask R-CNN では, 対象ごとにバイナリマスクを用意する必要がある.

果実ごとのバイナリマスクは, 図 2 に示すように, 果実画像を背景画像上に配置する際に生成する. 具体的には, 果実画像の不透明領域と透明領域とで区別してバイナリ画像を生成し, 全画素値を 0 で初期化した合成農園画像と同じサイズの画像上に配置する. バイナリ果実画像を配置する座標は, 果実画像を配置した座標と同一とする.

バイナリマスクは果実領域を表しているが, 生成した合成農園画像では, 果実画像を配置した後に葉画像や, ほかの果実画像が配置されるため, 果実部分には一部隠れた領域が生じる. このような手順によって合成農園画像とバイナリマスクを大量に生成することで, 隠れた果実の領域であっても元の領域を抽出可能な訓練データが得られる.

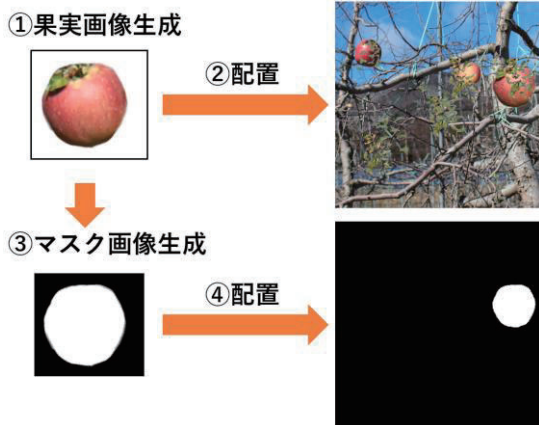


図 2: 果実ごとのバイナリマスクの生成

表 1: 果実検出精度の比較

	COCO	提案手法
AP	0.215	0.367
$AP^{IoU=.50}$	0.432	0.769
$AP^{IoU=.75}$	0.195	0.315
$AR^{max=100}$	0.384	0.492

3. 実験

提案手法による隠れ果実領域の学習の有効性を検証するために、比較実験を行った。比較対象は、COCO データセットを用いた学習済みモデル¹とした。この学習済みモデルは、80 種類のクラスに対応している。提案手法では、この COCO 学習済みモデルに対して、上記の手法で生成した隠れ果実領域を含む 3000 枚の訓練データを追加学習した。

学習モデルの検証用データとして、実際のリンゴ農園におけるモニタリング画像を 9 枚用い、そこから 90 枚に分割して切り出した画像を用いた。この 90 枚の画像に対して、リンゴの果実部分を手作業でアノテーション付けした。このとき、一部分が隠れている果実においては、作業者の判断で元の形を推定して領域を決定した。

果実の検出精度を確認するために COCO の精度検証手法である Average Precision (AP) と Average Recall (AR) [Lin 2014] を採用した。表 1 に各指標に対する果実の検出精度を示す。

4. 考察

4.1 隠れ果実領域の学習の有効性について

表 1 の結果より、各評価項目で提案手法のスコアが COCO 条件よりも上回っており、より正しく果実領域を抽出できていることがわかる。また、図 3 に示すように実際の検出領域を比較したところ、提案手法では果実の上に重なっている葉の影響を受けずに隠れた果実領域をより高精度に特定できていることがわかる。このように、提案手法によって隠れた果実領域の学習が可能であり、未学習の農園画像に対しても有効であることが示唆された。



図 3: 隠れ果実領域の検出例(緑: 真値, 赤: 推定値)

4.2 検証用のデータについて

隠れた果実領域は真値の取得が難しく、現状は作業員 1 名による手作業で対象領域を抽出し真値として扱っている。そのため、隠れた果実領域についてのデータの客観性が高いとは言い難く、より厳密な検証のためには、複数人での同じ対象の領域抽出を実施するのが望ましい。

5. まとめ

本研究では、果実の生長の把握のために、深層学習を用いて農園モニタリング画像から隠れた領域を含めた果実領域を抽出するための訓練データの生成方法を提案した。提案手法の有効性を検証するため、COCO 学習済みモデルに対して、生成した訓練データを学習させた際の果実検出精度の変化を評価した。今後、葉によって隠れている果実領域の抽出精度の向上や、検出した果実領域を用いた自動訓練システムなどの開発にも取り組む予定である。また、果実領域の正確な面積が求めれば、果実の生育情報抽出に有用であるため、正しい果実領域と予測した果実領域との面積比の平均値を求めることで、果実領域の面積抽出精度を評価する予定である。

参考文献

- [Bargoti 2017] Bargoti, S., Underwood, J.: Deep fruit detection in orchards. In 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 3626-3633 (2017)
- [Ren 2015] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks, Neural Information Processing Systems, Vol.1, pp.91-99, (2015)
- [Lin 2014] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L.: Microsoft COCO: common objects in context, In Proc. of the European Conference on Computer Vision, pp.740-755, (2014)
- [He 2017] He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN, In Proc. of the International Conference on Computer Vision, pp.1059-1067, (2017)

¹https://github.com/matterport/Mask_RCNN/releases/download/v2.0/mask_rcnn_coco.h5