

屋外環境下における深層学習を用いた果実検出手法の検討 A Study of Fruit Detection Method Using Deep Learning in Outdoor Environment

本間 大智[†]
Daichi Homma

馬淵 浩司[†]
Hiroshi Mabuchi

ゴウタム・チャクラボルティ[†]
Goutamu Chakraborty

松原 雅文[†]
Masafumi Matsuhara

1. はじめに

近年、農作業の機械化が進んでいるが、果樹栽培においては農業散布の機械化までに留まっており、農作業の大半を占める生育管理や収穫作業は人手により行われている。そこで、本研究では果樹栽培における農作業の効率化を目的とし、果実の成熟度を判別するシステムの構築を目指している。今回、その前段階となる果実の自動検出手法について検討する。

本研究では屋外環境下で撮影された画像を対象としているが、屋外環境下での物体検出は天候による照明変動により検出精度の低下が問題となる。そこで、照明条件の変動を考慮したデータ拡張を行い、YOLOv3 による検出実験を行う。検出結果をもとに適合率(Precision)、再現率(Recall)、F 値、平均適合率(Average Precision)を算出し、検出精度を検証する。

2. 関連研究

近年、深層学習分野の発展に伴い農業分野においても深層学習を用いた研究が行われている。果樹を対象とした農作物の検出では、Faster R-CNN を用いてマンゴーやアーモンド、リンゴ果実の検出を行う研究がある[1]。本研究では、枝や葉、果実同士によるオクルージョンが検出精度に影響を及ぼすと考え、オクルージョンに注視する。

3. 提案手法

本手法の流れを 3.1~3.3 に示す。

3.1 データ収集

果樹の撮影を行う際、雨天時を除き晴天、曇天時での撮影を行った。撮影期間は2019年10月から11月であり、総撮影枚数は992枚となる。各画像のサイズは4896×3672pixelである。撮影の際、果樹に対して3方向から撮影を行い、異なる照明条件の元データ収集を行った。

撮影した画像の内、537 枚については、枝や葉、果実同士の重なりによるオクルージョンの有無は考慮せず、全て1つのクラスとしてアノテーションを実施する。これらの画像を9:1の割合でランダムに選択し、484枚を学習に、53枚を検証用として用いる。オブジェクト数を表1に示す。

表 1 オブジェクトの数

	学習用	検証用
正解ラベル数	16923	1677

3.2 データ拡張

屋外環境下では天候や時間経過によって、太陽光の照射角度や強度が著しく変動する。従って、異なる撮影時間や天候下で撮影された画像から果実を検出するためには訓練画像の水増しを行い、学習モデルの汎化性を向上させる必要がある。

訓練画像に対してヒストグラム平坦化を含むコントラスト調整やガンマ補正による明るさの変更を行う。加えて、ガウシアンノイズやソルト&ペッパーノイズといったノイズ付与、ピントが合わずぼけた画像が撮影される可能性を考慮した平均化フィルタによるぼかしや、一般的なデータ拡張手法である回転処理といった前処理を行った。データ拡張により得られた画像の内訳は表2に示す。

表 2 データ拡張処理毎の画像枚数

元画像	明るさ	ノイズ	ぼかし	回転	合計
484	2456	986	484	1452	5324

3.3 物体検出部

本研究では撮影画像内の対象物の数が多いことから、その中でも検出精度及び速度共に優れているYOLOv3[2]を用いて果実検出を行った。

YOLOv3は入力画像をS×Sのグリッドセルで分割し、グリッドセル内に物体の中心がある場合そのグリッドセルが物体検出の対象領域となる。各グリッドセルでは事前に定義された複数の矩形を予測し、対応する信頼度の値を算出する。この値は、グリッドセル内に物体が含まれている可能性と、モデルによって予測された矩形の正確さを表している。信頼度は式(1)によって算出される。

$$IOU = \frac{\text{Area of Intersection}}{\text{Area of Union}} \quad (1)$$

IOU(Intersection over Union)とは、モデルによって予測された矩形と、正解となる矩形が重なり合う領域の面積に基づいて算出される。正解となる矩形と重なる面積が大きくなる程、信頼度の値は高くなる。

4. 実験

4.1 学習条件

YOLOv3 の学習条件を表3に示す。ネットワークへの入力サイズは416×416とし、元画像をリサイズして学習を行う。YOLOv3の学習はGPU(GeForce 1080 NVIDIA社製)を搭載したUbuntu18.04.3 LTS 上で行い、モデルの評価を行う際は最も予測精度の高いepochの重みを用いる。

[†] 岩手県立大学 Iwate Prefectural University

表 3 学習条件

Epoch 上限数	学習率	バッチ サイズ	ミニバッチ サイズ
10000	0.001	64	16

表 4 各モデルの検出結果

	Precision	Recall	F 値	AP ⁵⁰
データ拡張なし	83.1%	75.9%	79.3%	80.3%
データ拡張あり	78.3%	84.3%	81.2%	84.7%

表 5 検出結果の詳細

	TP	FP	FN
データ拡張なし	1274	259	403
データ拡張あり	1414	392	263

4.2 比較実験

照明条件の変動を考慮したデータ拡張による検出精度向上の有効性を確認するため、データ拡張を適用したモデルと適用していないモデルでそれぞれ果実検出を行い、精度の検証を行う。ここで、YOLOv3 による果実検出時の予測結果を判定する IoU 及び信頼度の閾値は共に 0.5 とする。

モデルの評価には、第 2.1 節で説明した 53 枚の検証用画像を用いた。各モデルの検出精度を評価指標となる Precision, Recall, F 値, Average Precision(AP)の算出結果を表 4 に、検出結果の詳細を表 5 に示す。TP(True Positive)はモデルが正しく対象物を検出した数、FP(False Positive)はモデルが誤って対象物と検出した数、FN(False Negative)はモデルが対象物を検出できなかった数を表している。正解データの例を図 1 に、データ拡張を行わないモデルの検出結果の例を図 2 に、データ拡張を行ったモデルの検出結果の例を図 3 に示す。図 2、図 3 から事前に付与したアノテーションの結果とおおむね一致した結果が得られており、果実のみを正しく検出していることが分かる。

4.3 考察

実験より、データ拡張を行った場合 Recall, F 値, AP⁵⁰ に関しては値の向上が確認された。正解データに対して検出漏れが少なく、予測精度の高いモデルを作成することができたと考えられる。唯一 Precision のみが低下しているが、表 5 より、TP が増加し FN が減少していることからデータ拡張によって検出性能が向上し全体的に検出数が増加したこと、実際には正しく検出されている矩形の予測領域に正解ラベルが存在しておらず、FP として扱われたことが原因として考えられる。

また、図 2 では果実を半分程度覆い隠すようなオクルージョンが発生している場合検出できなかったが、図 3 では正しく検出しており、オクルージョンが発生している果実に対するデータ拡張による有効性も示された。しかし、果実の輪郭が見えなくなるほど重度のオクルージョンが発生した場合、1 つの果実を複数の果実として誤検出している箇所も見られる。

従って、学習データの見直しに加えラベルを付与する際にはオクルージョンの有無によってクラスを設定し、再度検出性能の確認を行う必要がある。



図 1: 正解データ



図 2: 検出結果(データ拡張なし)

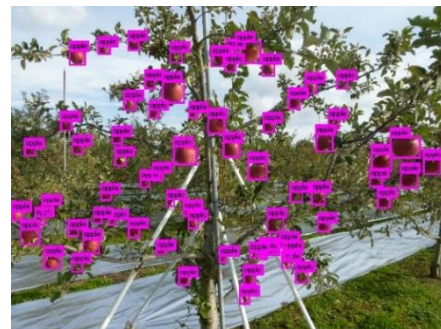


図 3: 検出結果(データ拡張なし)

5. おわりに

今回、屋外環境下で撮影された画像に対して照明条件の変動を考慮したデータ拡張を行い、果実検出手法の検討を行った。実験では、データ拡張の有無で検出精度の比較検証を行った。その結果、検出精度の向上が認められデータ拡張の有効性を示すことができた。

今後の課題として、果実検出における誤検出の原因となるオクルージョンへの対応が挙げられる。本研究で使用した学習データはオクルージョンの有無を考慮せずラベルを付与しているが、オクルージョンの有無を考慮した 2 クラスでのラベル付与を検討する予定である。

謝辞

本研究の一部は JSPS 科研費 18K11358 の助成を受けたものである。

参考文献

- [1] S. Bargoti, J. Underwood, "Deep fruit detection in orchards", Aust. Centre Field Robotics 1-8, (2016).
- [2] J. Redmon, A. Farhadi, "Yolov3 An incremental improvement", arXiv preprint arXiv:1804.02767, (2018).