

セグメンテーションラベルを用いたユズ果実の物体認識モデルの精度向上

1275104 杉山 理裕 【 知能情報学研究室 】

Performance Improvement of Object Recognition Models for Yuzu Fruit Using Segmentation Labeling

1275104 SUGIYAMA, Takahiro 【 Intelligent Informatics Lab. 】

1 はじめに

農業分野において、収量予測の精度を高めるためには、果実数の計測は重要な課題であり、AI（人工知能）による画像物体認識を用いた手法が近年注目されている。

従来の果実数計測では、物体認識モデルである YOLO を用いる手法が一般的である。この手法では、果実に対して手動でラベル付けを行い、モデルを学習させる必要がある。しかし、膨大な数のラベル付けを手動で行うためには多大な時間やコストがかかるという問題がある。

本研究では、果実検出の効率化と精度向上を通じて、従来の手動ラベル付けによる課題を解決するため、セグメンテーションのための基盤モデルである SAM(Segment Anything Model)[1] を用いたラベル付けの自動化を提案する。これにより、時間やコストの削減を図り、従来の手法との精度比較を行う。さらに、自動生成したラベルと手動生成したラベルの品質を比較評価するためにラベルの信頼性を評価するライブラリである Cleanlab[2] を活用して評価する。

2 提案手法

2.1 SAM によるセグメンテーション

SAM を活用したセグメンテーション手法による物体検出モデルの精度向上を提案する。SAM は、入力されたバウンディングボックス内で物体と背景を精密に区別し、物体の輪郭に沿ったセグメンテーションマスクを生成する。具体的には、バウンディングボックス内の各ピクセルについて、「物体」か「背景」かを学習済みのパターンに基づいて予測し、物体の境界線や色調、質感の違いを活用して関連するピクセルを選択する。その結果、物体に対応するピクセルのみを抽出し、背景を除去したマスクを生成する。

SAM を用いることで、物体が背景や他の物体に隠れている場合でも、物体の輪郭を正確に切り取れるため、ノイズの影響を低減できるという利点がある。本研究では、YOLOv8 の学習済みモデルの検出結果と SAM のセグメンテーション結果を利用し精度向上を図る。図 1 は SAM を用いた自動アノテーションによる教師データ増加手法の概要図である。

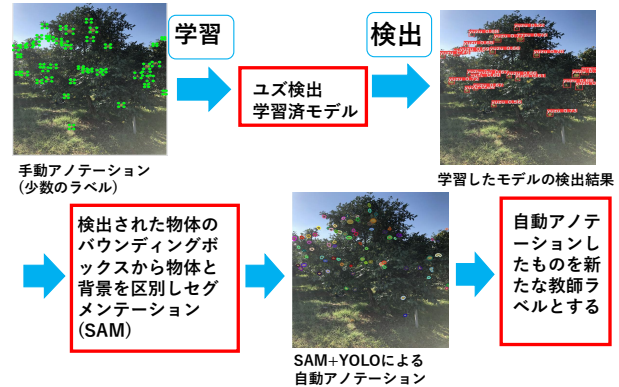


図 1: SAM を用いた自動アノテーションによる教師データ増加手法の概要図

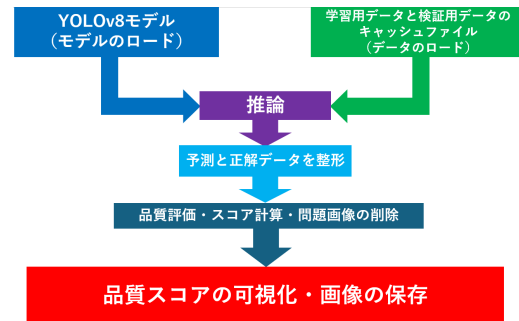


図 2: ラベル品質評価のシステム概要図

2.2 Cleanlab によるデータクレンジング

手動でアノテーションしたラベルの品質と自動アノテーションしたラベルの品質を比較するため、Cleanlab を利用し、品質スコア条件以下のものの画像の除去をする。具体的には、学習済みの YOLOv8 モデルを用いて学習用および検証用データに対して推論を実施し、その結果得られた予測ラベルと正解ラベルを Cleanlab 用に整形することで品質スコアを計算し、設定した閾値未満の画像の削除と閾値以上の画像の保存を行う。基準を満たす画像のみを保存することで、データセットの全体的な品質を向上することができる。図 2 はラベル品質評価のシステム概要図である。

3 実験方法

3.1 SAM による自動アノテーション

ユズの実を対象に、手動で作成したアノテーションと自動化手法を組み合わせた物体検出・セグメンテーションの精度向上を検証する。まず、手動で LabelImg を使用し、ユズの実に対してバウンディングボックスを作成したアノテーションを行い、YOLOv8n で学習を実施する。データセットは、学習用 238 枚、検証用 108 枚、テスト用 68 枚の画像を使用した。次に、SAM を用いて、YOLOv8n の物体検出結果で得られたバウンディングボックスを基に、詳細なセグメンテーション解析を行い、バウンディングボックス内のピクセルごとに物体と背景を区別し、図 1 のように物体の形状に沿ったセグメンテーションマスクを生成した。生成したマスクの結果を基にアノテーションファイルを作成し、再び YOLOv8n に学習させることで、自動化された学習手順の評価を行う。また、手動アノテーションモデルと比較して、自動化手法が精度に与える影響を検証する。さらに、手動で作成したアノテーションの画像やラベルの数をランダムに 1/2, 1/3... と減少させた場合、どの程度精度を維持できるかも評価する。

3.2 ラベル品質評価

手動と自動で付けたラベルに対して、どれだけアノテーション品質が高いか検証する。方法としては学習用データ 238 枚と検証用データ 108 枚に対して、Cleanlab を用いたデータクレンジングを実行し品質スコアの低いものを Cleanlab によって除去する。品質スコアが 0.3 以上のものと 0.5 以上のもので分類しどれだけの枚数が出るか検証する。

4 結果・考察

手動で付けたラベルを学習させたモデルの AP は 0.802, F1 スコアは 0.75 であった。一方、自動化による学習を 15 回繰り返した結果、7 回目の学習モデルが最も良く、AP 値が 0.98, F1 スコアが 0.93 を達成した。テスト用画像での比較でも、自動化モデルが手動モデルを上回る結果を示した。さらに、画像またはラベルをランダムに削減した場合、画像を 1/10 に減らしても精度は一定程度保証されたが、ラベルを 1/2 減少させた時点で、手動モデルよりも精度が劣る結果となった。このことから、適切にアノテーションされた学習・検証データが、自動化モデルの性能向上に不可欠であると考えられる。

また、Cleanlab によるデータクレンジングの結果、手動モデルでは 0.3 以上の信頼度を持つデータが学習用で 211 枚、検証用で 90 枚、0.5 以上がそれぞれ 139 枚と 48 枚であった。一方、自動化モデルでは、0.3 以上が学習用で 217 枚、検証用で 101 枚、0.5 以上がそれぞれ 169 枚と 69 枚となり、自動化によってアノテーションの品質が向上していることが示された。

表 1: 手動学習と自動学習の比較

学習タイプ	AP 値	F1 スコア
手動学習	0.802	0.75
自動学習 (最良回数:7 回)	0.98	0.93

表 2: 画像およびラベル減少の比較

条件	結果
画像を 1/10 減少	精度を保証
ラベルを 1/2 減少	手動より劣る

表 3: Cleanlab によるデータクレンジングの結果

学習タイプ	品質スコア条件	学習枚数	検証枚数
手動学習	0.3 以上	211 枚	90 枚
	0.5 以上	139 枚	48 枚
自動学習	0.3 以上	217 枚	101 枚
	0.5 以上	169 枚	69 枚

5 おわりに

本研究では、手動によるアノテーションと、SAM を用いた自動化アノテーションの精度比較を行い、さらに Cleanlab を用いた品質評価によって両手法のアノテーション品質の向上を検証した。その結果、自動化によるアノテーションは、手動のアノテーションに比べて精度および効率の面で優れていることが示された。特に、自動化を繰り返し実行することで検出精度が向上し、最も良い結果では AP 値が 0.98, F1 スコアが 0.93 を達成した。また、Cleanlab を用いたデータクレンジングにおいても、自動化されたアノテーションデータの方が品質スコアが高いデータの割合が多いことが確認された。

これらの結果から、自動化によるアノテーションの導入は、時間やコストの削減だけでなく、検出精度やアノテーション品質の向上にも有効であると結論づけられる。また、ランダムに学習データやラベルの一部を削減した場合でも、一定の範囲内で精度が保証されることが示された。

参考文献

- [1] Kirillov, *et al*, "Segment Anything", Meta AI Research, FAIR, 2023
- [2] Curtis, *et al*, "Confident Learning: Estimating Uncertainty in Dataset Labels", Journal of Artificial Intelligence Research (JAIR), 2020