

圃場の特性を利用した画像処理による株の検出と特徴量を拡張したスタッキングによるキャベツの葉数推定

～ 前提条件の設定により少ない学習データでも安定して高精度に推定可能なアルゴリズム ～

中西 凜太郎†, 王 逸倫†, 関 峻太†, 福井 雅也†

†: みずほリサーチ&テクノロジーズ株式会社,
{ rintaro.nakanishi, yilun.wang, ryota.seki, masaya.fukui }@mizuho-rt.co.jp

本検討では、圃場画像特有の傾向を、株の検出用アルゴリズム構築および葉数推定のための特徴量作成に活用した。各画像のベースモデルの確信度として、ベースモデルの推定値の分散を特徴量に加えたことで、少ない学習データでも安定して高精度な葉数推定を実現した。

1. はじめに

本課題では、複数株のキャベツが写った圃場の画像から、株の有無と各株の葉数を推定するモデルを構築する。モデル構築に使用可能な画像枚数は8枚と限られており、バリエーションの少ない学習データで安定した推定を実現することが求められる。

我々は、少ない学習データにおいても、推定の安定性と精度を向上させるため、圃場の画像特有の傾向を基に5つの条件を設定した。この条件に基づき、株の位置は画像処理を用いて決定論的方法により検出し、葉数は画像および3種類のベースモデルから算出される特徴量をもとに推定した。

2. 前提条件の設定

我々は、同一画像内では株の配置や大きさについて共通する特徴があると考え、この特徴が非公開の採点対象画像でも共通するものと仮定した。抽出した特徴を以下 i ~ v に示す。

- i. 株は縦2列に並ぶ(以下、各列を畝という)
- ii. 株は上下に概ね一定間隔で並ぶ
- iii. 株の幅と高さは概ね均一
- iv. 株の大きさと葉数は正の相関を持つ
- v. 同一画像内の株の葉数は概ね均一

3. アルゴリズム概要

本手法における、各株の検出および葉数の推定アルゴリズムを以下にそれぞれ説明する。

3-1. 株の検出

株の検出手順は、以下1~6の通り。

1. HSV色空間(h, s, v)で画像から緑色($70^\circ \leq h \leq 170^\circ$, $15\% \leq s \leq 100\%$, $15\% \leq v \leq 100\%$)の画素を抽出
2. 前提条件 i に基づき、1で抽出した画素をx軸(横軸)に射影して得られるヒストグラムの途切れ目を基準として、畝のx軸の範囲を2つ抽出(図1下部)
3. 前提条件 ii に基づき、各畝において、1で抽出した画素をy軸(縦軸)に射影して得られるヒストグラムの局所最小値を基準として、各株のy軸の範囲を抽出(図1左右)
4. 前提条件 iii に基づき、2, 3で得られた各株のx方向の長さ(幅)とy方向の長さ(高さ)が概ね同じ値になるよう、幅*0.8 ≤ 高さ ≤ 幅*1.2の範囲に高さを調整
5. 4で得られたy軸の範囲ごとに、株全体が写る最小サイズまで、x軸の範囲を調整
6. 4, 5で得られた各軸の範囲内を株とする



図1 株の検出結果例

3-2. 各株の葉数の推定

葉数の推定は、以下A) ~C) に示す3種類

のベースモデルと圃場画像から算出される特徴量をもとに推定した。

- A) ResNet[1]で抽出した特徴量を用いて葉数を推定する線形回帰モデル
- B) Mask R-CNN[2]で推定したセグメント個数を葉数とするモデル
- C) RTMDet[3]で推定したセグメント個数を葉数とするモデル

最終的な葉数は、以下 1~3 を結合した 7 次元の特徴量を入力とする、ランダムフォレスト回帰（以下 RF という）で推定した。

1. 株の範囲の面積 S_i
2. A) ~C) で推定した各株の葉数 N_i^α
3. 式 (1) にて算出する、A) ~C) における、同一画像内の葉数の推定値の分散

$$\sigma^\alpha = \frac{1}{M} \sum_{i=1}^M (N_i^\alpha - \overline{N^\alpha})^2 \quad (1)$$

ここで、 M は画像内の株の総数、 N は葉数の推定値、 i は株のインデックス ($i=1, 2, \dots, M$)、 α はベースモデル A) ~C) である。通常のスッキングでは N_i^α のみの特徴量とするが、本手法では S_i と σ^α を特徴量に追加した。 S_i を追加した理由は、前提条件 iv より、株の面積が葉数の推定に有用であると考えたためである。 σ^α を追加した理由は、前提条件 v より、精度の高いモデルでは σ^α が小さくなる傾向があると考えたためである。すなわち、 σ^α を各ベースモデルの推定結果に対する確信度とみなし、RF がどのベースモデルに重きを置くべきかを、入力画像ごとに動的に変化させることを狙いとしている。

4. 結果・考察

本課題では、4 つの生育段階ごとに、各 2 枚の圃場の画像（合計 8 枚）が与えられる。本手法では、学習に全ての生育段階の画像を各 1 枚用い、残りを評価に用いた。

株の検出では、評価用画像において、検出すべき株を全て検出できた。

各株の葉数の推定について、評価画像に対する MAE を表 1 に示す。

表 1 評価画像に対する MAE

	A	B	C	stacking	本手法
誤差 [枚/株]	0.96	1.96	2.00	1.04	0.92

※ 本手法の処理時間: 35[秒/画像]

※ CPU: 13th Gen Intel(R) Core(TM) i5-1350P 1.90GHz,

Memory: 32GB, OS: Windows 11 Enterprise

ここで、ベースモデル A、B、C の結果をそれぞれ”A”、”B”、”C”、 N_i^α のみの特徴量とする通常のスッキングモデルの結果を”stacking”、本手法の結果を”本手法”として示す。

表 1 から、本手法はいずれのベースモデルよりも高い精度で各株の葉数を推定出来ていることが分かる。画像から葉数を直接推定するベースモデル A と、セグメンテーション結果から間接的に推定するベースモデル B、C の間で、予測精度が高くなる画像の傾向に違いがあったことで、誤差が相殺されたと考えられる。

また、通常のスッキングモデルと本手法の比較では、本手法の方が葉数の推定精度が高かった。従って、特徴量として加えた S_i と σ^α の情報が、本手法の精度向上につながったことが確認できた。

5. まとめ

本手法では、限られた枚数の圃場画像を用いて、キャベツの株の有無および各株の葉数を精度高く推定するモデルを構築した。評価の結果、株の位置は全ての評価画像において正しく検出できた。また、葉数の推定精度は、各ベースモデル単体およびスッキングモデルを上回る結果が得られ、特徴量に株の大きさや推定値の分散を追加することが、精度向上に寄与することを確認した。以上より本手法は、学習データが少ない状況においても、圃場画像の特性を活かすことで、株の検出および葉数の推定を高い精度で実現できることを示した。

参考文献

- [1] Kaiming H. et.al, “Deep Residual Learning for Image Recognition,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778
- [2] Kaiming H. et.al, “Mask R-CNN,” 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980-2988.
- [3] Chengqi L. et.al, “RTMDet: An Empirical Study of Designing Real-Time Object Detectors,” arXiv preprint arXiv:2212.07784, 2022.