

## UAV 空撮画像と深層学習を用いた高知県産カンキツ類の生育計測と収量予測

石橋 璃貴 【 知能情報学研究室 】

## Study on Growth Measurement and Yield Prediction of Citrus Fruits in Kochi Using UAV Imagery and Deep Learning

ISHIBASHI, Riki 【 Intelligent Informatics Lab. 】

## 1 はじめに

本県の特産である土佐文旦などのカンキツ栽培において、着花量等の生理生態情報の把握は早期の出荷計画や摘果などに必要な人的資源を行う上で重要な工程であると考えられるが、既存手法のサンプリング法では、多大な労力が課題となっている。この課題の解決手法として UAV 空撮画像を色閾値2値化処理を行うことで着花量の定量的な評価を行う手法が提案された [1]。しかし、この手法には、手作業での樹冠領域を抽出のコストや、RGB の固定値で2値化を行っているため、ロバスト性に問題があると考えられる。そこで本研究では YOLO, SAM で樹冠領域の抽出を自動化し、U-Net でロバスト性の高い着花量の定量化手法を提案する。

## 2 提案手法

本研究では、UAV 空撮画像から土佐文旦の収量を予測するため、画像中央の樹体領域と花領域の抽出を階層的な手法を提案する。提案手法の全体図を図1に示す。処理は以下の3段階で構成する。

1. 樹体検出と ROI (Region of Interest) 推定：  
YOLOv11n による BBOX での樹体検出
2. 樹体のセグメンテーション：BBOX を入力として受け取り SAM による樹体領域のセグメンテーション
3. 花領域の推定：U-Net による花房のセマンティックセグメンテーション

この処理により、地面の雑草や隣接する樹木などのノイズを自動で排除し、対象樹木の花のみを解析することを可能にする。

## 3 実験

本研究では、高知県農業技術センター果樹試験場内露地圃場で UAV により、地上高 7.5~9.0m から撮影されたユズ樹体と土佐文旦樹体を対象とした空撮画像を用いる。画像サイズは 4000 × 2250px である。

## 3.1 樹体認識モデル

樹体検出モデルでは、空撮画像全 226 枚を訓練データ：173 枚と検証データ：53 枚に 8:2 で分割する。教師

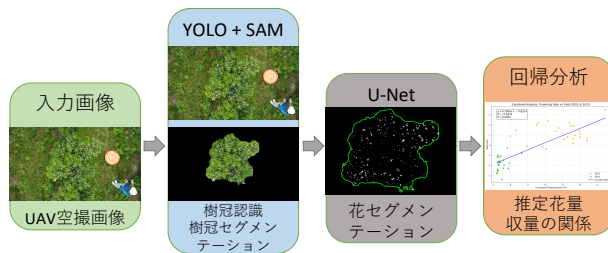


図 1: 提案手法の全体概要

ラベルとして画面中央の樹体のみにもアノテーションを行い、位置情報と物体クラスのラベルを生成する。認識モデルとして YOLOv11n に転移学習を用いて、入力画像サイズは 1280px, epochs は 200 として学習を行う。樹体認識精度の評価指標として Precision と Recall, 正解領域と予測領域の重なりが 50 %以上のものを正解とした Average Precision@50(AP@50) を用いる。

## 3.2 花領域セグメンテーションモデル

花領域セグメンテーションモデルでは、従来手法で性能が良かった RGB の閾値2値化処理と U-Net でのセグメンテーション結果を比較する。使用するデータセットは、UAV 空撮画像から中央の 1200px を抽出した画像を 150px に分割し、手作業で花を塗りつぶした正解マスク画像と RGB 画像の組み合わせ全 1179 枚作成する。この画像組 8 : 2 で分割し、訓練データ：943 枚とテスト：236 枚とする。また t 訓練データは 5 分割交差検証を行い、モデルが 2 値化マスク画像を生成する際の花とする閾値を探索する。その結果から閾値 0.8 が最も性能が良かった。この閾値を用いてテストデータ上で性能評価を行う。評価指標として、予測領域と正解領域の重なり割合である mIoU(mean Intersection over Union) と正解領域の画素と予測領域画素の線形回帰の決定係数とモデルの検証データ上での適合度と複雑さのバランスを定量化する AIC (赤池情報量基準) を用いて比較する。

## 3.3 収量予測モデル

YOLO + SAM で抽出した樹冠領域内で花領域セグメンテーションモデルで予測花領域を求め、各樹体から樹

表 1: 樹体検出モデル (YOLOv11n) の評価結果

Precision	Recall	AP@0.5
0.997	1.000	0.995

表 2: 定量的評価結果の比較

Method	$R^2$	AIC	mIoU
Proposed	<b>0.99</b>	<b>3437</b>	<b>0.72</b>
RGB Threshold	0.79	3551	0.33

冠領域あたりの花領域を計算し、推定花量を算出する。推定花量を説明変数、1 樹体あたりの収穫量 ( $kg/m^3$ ) を目的変数として線形回帰、二次多項式回帰、SVR を用いて分析を行う。評価指標として自由度調整済み決定係数 *Adjusted R<sup>2</sup>* を使用する。

## 4 実験結果・考察

### 4.1 樹体認識モデル

表 1 に YOLOv11n による樹体検出の評価結果を示す。Precision は 0.997, Recall は 1.000, AP@0.5 は 0.995 といずれも極めて高い値を示した。特に Recall が 1.000 であることから、検証データにおいて、対象となる樹体を見逃すことなく検出することに成功している。この高い検出精度により、本モデルは後段の SAM によるセグメンテーション処理に対して正確な ROI を推定可能であることが示された。

### 4.2 花領域セグメンテーションモデル

表 2 に、テストデータにおける定量的評価の結果を示す。また、結果画像を図 2 に示す。

比較の結果、提案手法である U-Net は、すべての評価指標において従来手法を大きく上回る性能を示した。まず、領域の重なり具合を示す mIoU において、RGB 閾値法は 0.3285 と低い値にとどまった。これは、閾値処理が色情報のみに依存するため、背景ノイズ等や日光の照り返しを花として誤検出したこと、日陰にある花が検出困難だった事が原因と考えられる。一方、U-Net は 0.72 と高い精度を達成しており、ノイズの影響を受けずに花の特徴を学習可能であることが示唆された。

また、正解画素数と予測画素数の関係を示す *adjustR<sup>2</sup>* においても、U-Net は 0.99 という高い相関を示した。さらに、AIC においても、提案手法は従来手法より低い値 (3437) となり、統計的に有意に優れたモデルであると判断できる。以上の結果から、U-Net による花領域抽出は、収量予測を行うための信頼性の高い手法であるといえる。

### 4.3 収量予測モデル

表 3 に収量予測精度の比較を示す。結果として、提案手法と SVR の組み合わせが最高精度 ( $R^2 = 0.714$ ) を

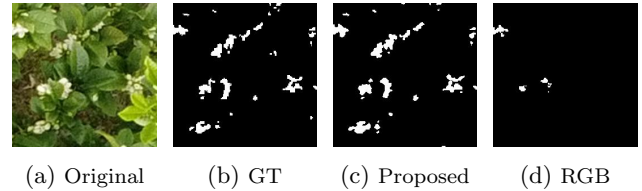


図 2: 花領域推定の結果比較

表 3: 収量予測における各回帰モデルの *adjR<sup>2</sup>* の比較

回帰モデル (Model)	RGB	U-Net (Proposed)
Linear Regression	0.623	<b>0.641</b>
Polynomial (2nd)	<b>0.679</b>	0.676
SVR (Non-linear)	0.707	<b>0.714</b>

記録した。モデル間の比較では、線形回帰 ( $R^2 \approx 0.63$ ) に対し、多項式回帰や SVR といった非線形モデルが優位性を示した。これは、着花過多による養分競合や摘果の影響で生じる収量の飽和を、これらのモデルが適切に捉えたためと考えられる。また、画像処理手法においては、提案手法が RGB 法を上回った。RGB 法はノイズや落花の誤検出を含んでいたのに対し、提案手法は U-Net による高精度な抽出により、ノイズの少ない着花量を算出できたことが精度向上 ( $0.707 \rightarrow 0.714$ ) に寄与した。結論として、UAV 空撮画像からの収量予測には、正確な花領域抽出と非線形回帰モデルの統合が有効であることが示唆された。

## 5 おわりに

本研究では、UAV 画像を用いた土佐文旦の収量予測に対し、YOLOv11・SAM・U-Net を統合した手法を提案した。実験の結果、U-Net による花抽出は IoU 0.99 の高精度を達成した。また、SVR による予測モデルは決定係数 0.714 を記録し、推定花量と収量の非線形性を考慮する有効性を実証した。今後は複数年データの検証や、他のカンキツ品種への適用による汎用性評価と収量予測モデルの定量評価を目指す。

## 参考文献

- [1] 谷本他, “露地土佐文旦における着花量の簡易測定技術および葉面積指数の測定技術,” 農業情報研究, Vol. 33, No. 4, pp. 131–143, 2025.
- [2] O. Ronneberger et al., “U-net: Convolutional networks for biomedical image segmentation,” Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234–241, 2015.
- [3] N. Ravi, et al., “SAM 2: Segment Anything in Images and Videos,” arXiv preprint arXiv:2408.00714, 2024.