

令和7年度  
修士学位論文

# 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

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

石橋 璃貴

指導教員 吉田 真一教授

2026年2月27日

高知工科大学大学院 工学研究科 基盤工学専攻  
情報学コース

# 要旨

## 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

国内の農業従事者の減少に伴い、カンキツ栽培においても省力化と生産安定化に向けたスマート農業技術の実装が急務である。特に隔年結果性の強いユズや土佐文旦において、安定生産を実現するためには、着花量の早期把握に基づく適正な着果管理が不可欠である。露地土佐文旦における着花量の測定に関しては、谷本らにより UAV 空撮画像の RGB 色閾値処理による 2 値化画像解析を用いた簡易測定技術が提案されている。同手法は現場利用の可能性を示したが、複雑な背景を有する露地環境下において、手作業で樹冠領域の抽出、および花量推定時のロバスト性に課題が残ると考える。そこで本研究では、これらの課題を解決し、着花量を高精度かつ全自動で定量化する手法提案する。提案手法は以下の 3 段階で構成される。第一に、物体検出モデル YOLOv11 を用いて広域空撮画像から樹体位置を特定する。第二に、基盤モデル SAM 2.1 (Segment Anything Model 2.1) を導入し、樹冠領域のみを高精度にセグメンテーションすることで、谷本らの手法で課題であった背景ノイズの影響を排除する。第三に、抽出された樹冠画像に対し U-Net を適用することで、花領域の画素単位での検出を実現する。実証実験の結果、本手法は従来手法と比較して検出精度が大幅に向上し、複雑な背景下でもロバストな定量化が可能であることが確認された。本研究の成果は、露地果樹における生理生態情報の見える化を高度化し、データ駆動型の栽培管理体系の構築に寄与するものであると考える。

**キーワード** IoP, U-Net, YOLO, SAM

# Abstract

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

With the declining number of agricultural workers in Japan, the implementation of smart agriculture technologies for labor-saving and stable production has become an urgent task in citrus cultivation. Particularly for Yuzu and Tosa Buntan, which exhibit strong alternate bearing habits, appropriate fruit management based on an early assessment of flowering intensity is essential for achieving stable production. Regarding the measurement of flowering intensity in open-field Tosa Buntan, Tanimoto et al. [14] proposed a simplified measurement technique using binary image analysis based on RGB color thresholding of UAV aerial images. Although this method demonstrated potential for field application, challenges remain regarding the robustness of flower estimation in complex open-field environments, specifically due to the need for manual extraction of the tree canopy. To address these issues, this study proposes a fully automated and high-precision method for quantifying flowering intensity. The proposed framework consists of three stages: First, the object detection model YOLOv11 is employed to identify tree positions from wide-area aerial images. Second, the foundation model SAM 2.1 (Segment Anything Model 2.1) is introduced to precisely segment only the tree canopy regions, thereby eliminating the influence of background noise, which was a limitation in Tanimoto et al.'s method. Third, U-Net is applied to the extracted canopy images to

achieve pixel-level detection of flower regions. Experimental results confirmed that the proposed method significantly improved detection accuracy compared to conventional methods and enabled robust quantification even under complex background conditions. The findings of this study contribute to advancing the visualization of physiological and ecological information in open-field fruit trees and establishing a data-driven cultivation management system.

***key words*** IoP, U-Net, YOLO, SAM

# 目次

<b>第 1 章</b>	<b>序論</b>	<b>1</b>
<b>第 2 章</b>	<b>関連研究</b>	<b>3</b>
2.1	IoP (Internet of Plants) プロジェクトにおける生理生態情報の可視化 . . .	3
2.2	露地果樹栽培における課題と画像解析手法 . . . . .	4
2.3	物体認識・セグメンテーション技術 . . . . .	5
2.3.1	CNN (Convolutional Neural Networks) . . . . .	5
2.3.2	YOLO (You Only Look Once) . . . . .	5
2.3.3	SAM (Segment Anything Model) . . . . .	6
2.3.4	U-Net . . . . .	7
<b>第 3 章</b>	<b>提案手法</b>	<b>9</b>
3.1	BBOX での樹冠領域推定 . . . . .	10
	YOLO による処理過程 . . . . .	11
	Non-Maximum Suppression (NMS) . . . . .	12
	空間的ヒューリスティクスによる対象選定 . . . . .	12
3.2	SAM による樹冠領域のセマンティックセグメンテーション . . . . .	12
3.2.1	BBOX 位置情報プロンプトベースのマスク生成 . . . . .	13
3.2.2	マスクの統合と精緻化 . . . . .	13
3.3	U-Net による花領域 Semantic Segmentation . . . . .	14
3.3.1	ネットワークアーキテクチャ . . . . .	14
3.3.2	損失関数の設計 . . . . .	16
3.3.3	スライディングウィンドウによるパッチ推論 . . . . .	17
3.3.4	最終的な花領域の決定 . . . . .	17

## 目次

3.3.5	着花率の定義と定量的評価 . . . . .	18
3.4	推定花量に基づく収量予測モデルの構築 . . . . .	19
3.4.1	線形回帰モデル (Linear Regression) . . . . .	20
3.4.2	2次多項式回帰モデル (Polynomial Regression) . . . . .	20
3.4.3	サポートベクター回帰 (Support Vector Regression) . . . . .	21
<b>第4章</b>	<b>実験内容</b>	<b>22</b>
4.1	樹冠領域推定モデルの評価実験 . . . . .	22
4.1.1	モデル学習条件とハイパーパラメータ . . . . .	23
4.1.2	推論設定 . . . . .	24
4.1.3	評価指標 . . . . .	25
4.2	花領域検出モデルの評価実験 . . . . .	25
4.2.1	データセットの構築 . . . . .	25
4.2.2	比較・評価手法 . . . . .	26
	提案手法: U-Net アンサンブル . . . . .	26
	比較手法 1: NTSC 加重グレースケール法適用後, RGB 色閾値 2 値 化処理 . . . . .	27
	比較手法 2: RGB 閾値法 . . . . .	27
4.2.3	評価指標 . . . . .	27
	Intersection over Union (IoU) . . . . .	28
	面積相関解析 (Regression Analysis) . . . . .	28
4.2.4	実験条件の実装詳細 . . . . .	29
4.3	推定花量に基づく収量予測モデルの評価実験 . . . . .	29
4.3.1	データセットの前処理と補正 . . . . .	29
4.3.2	回帰モデルの選定 . . . . .	29
4.3.3	評価指標 . . . . .	31

## 目次

<b>第 5 章</b>	<b>実験結果と考察</b>	<b>32</b>
5.1	樹冠領域推定モデルの評価実験結果 . . . . .	32
5.1.1	考察 . . . . .	32
5.2	花領域検出モデルの評価実験結果 . . . . .	33
5.2.1	考察 . . . . .	34
5.3	推定花量に基づく収量予測モデルの評価実験結果 . . . . .	37
5.3.1	収量予測精度の定量的評価 . . . . .	37
5.3.2	回帰分析による傾向の可視化 . . . . .	37
5.4	考察 . . . . .	39
5.4.1	画像解析手法の比較と精度の要因 . . . . .	39
5.4.2	花量と収量の非線形性に関する考察 . . . . .	39
5.4.3	推定花量が収量モデルに及ぼす影響 . . . . .	40
	標準化データを用いた解析結果 . . . . .	40
	収量決定要因における環境と管理の支配性 . . . . .	41
5.4.4	本システムの有用性の再定義 . . . . .	42
<b>第 6 章</b>	<b>結論</b>	<b>43</b>
	謝辞	45
	参考文献	46

# 目次

3.1	提案手法の概要	10
3.2	樹冠領域 BBOX から生成したセグメンテーション画像	15
3.3	U-Net 構成図	16
3.4	花領域セグメンテーション結果画像	19
3.5	推定花量の可視化画像	20
4.1	樹冠領域推定モデルの学習データセット	23
4.2	IoU の計算式	28
5.1	樹冠領域推定モデル (YOLOv11nano) の出力画像	33
5.2	U-Net(a) と従来手法 (b)-(d) の散布図	35
5.3	提案手法と従来手法の比較結果	36
5.4	手法ごとの推定花量と実収量の回帰分析結果	38
5.5	推定花量及び収量を年ごとに標準化した際の回帰分析結果	41

# 表目次

4.1	Hyperparameters used for training YOLO11n . . . . .	24
4.2	2022 年および 2023 年の収量データ (単位: kg/m <sup>3</sup> ) . . . . .	30
4.3	着花率の補正係数 . . . . .	30
5.1	樹冠領域推定モデル (best.pt) の評価結果 . . . . .	32
5.2	Comparison of segmentation performance (IoU) . . . . .	34
5.3	各手法および回帰モデルにおける自由度調整済み決定係数 ( $R^2$ ) の比較 . . .	37

# 第 1 章

## 序論

日本の農業における基幹的農業従事者数は年々減少している。農林水産省の統計によると、基幹的農業従事者数は平成 27 年（2015 年）の 175 万 7 千人から、令和 7 年（2025 年）の概数値では 102 万 1 千人へと、この 10 年間で約 4 割以上減少している [1]。また、その内訳を見ても、令和 7 年時点で 65 歳以上の従事者が 71 万人（全体の約 7 割）を占めており、平均年齢は 67.6 歳と依然として高い水準にある。このように、従事者の高齢化および後継者不足による人的リソースの枯渇は、国内の生産基盤を急速に弱体化させている。このような情勢の中、農業の持続的な発展のためには、限られた労働力で生産性を維持・向上させるための、スマート農業技術の更なる実用化と現場実装が強く求められている [2]。スマート農業とは、ロボット、AI (Artificial Intelligence)、IoT (Internet of Things) 等の先端技術を活用する農業のことであり、現在、様々な部門で実装が進んでいる。特に高知県では、作物生産を決定づける光合成や成長等の生理生態情報を「見える化」し、それに基づく合理的な営農支援情報として「使える化」を行い、産地全体で情報を「共有化」する仕組みである「IoP (Internet of Plants)」[3] の研究および実装が、施設園芸を中心に進められている。しかし、施設園芸と比較して環境制御が困難な露地果樹栽培においては、広大な圃場におけるデータ収集の難しさや、気象条件等の外的要因によるノイズの多さから、IoP の実装は遅れているのが現状である。地域特性に応じた多種多様な果樹品目においてデータを効率的に活用するためには、品目ごとの生理生態情報を野外環境下で正確に「見える化」する技術の開発が不可欠である。果樹での「見える化」が可能な生理生態情報として、着花数、着果数、葉面積指数 (Leaf Area Index: LAI)、樹勢等が挙げられる。特に、本研究が対象とするユズや土佐文旦などのカンキツ類は一般的に隔年結果性が強く、豊作年と不作

年を繰り返すことで果実収量、収益が不安定となりやすい。出荷量を早期に予測し、適切な摘果や施肥計画を立てることは、経営の安定化だけでなく、収穫や選果に必要な人員配置などの労務管理においても重要である。しかし、従来の着花量調査は枝先 20cm 調査法 [4] のような部分的なサンプリング手法に依存しており、広範囲な果樹園全体を把握するには多大な労力を要する上、樹体ごとの個体差を正確に反映できないという課題があった。そこで近年、UAV（ドローン）による空撮画像と深層学習 (Deep Learning) を用いた広域かつ高精度なモニタリング技術が注目されている [8]。画像処理による花の定量化に関する研究はリンゴやマンゴー等で報告されているが [10, 9]、背景が複雑な露地環境下において、個々の樹体を正確に認識し、かつ微細な花を画素単位で抽出することは依然として困難な課題である。単に画像全体から花を検出するだけでは、隣接する樹木や雑草との区別がつかず、樹体ごとの生育情報の管理には繋がらない。この課題を解決するため、露地栽培カンキツ類を対象した UAV 空撮画像の着花量推定手法が提案された [14]。この手法では、手作業で樹冠領域の抽出しその領域を RGB 閾値処理を行うことでバイナリマスクを生成するものであった。この手法により、画素単位で花を選別でき、推定花量として空撮画像内の花量を定量的な評価可能にした。また、収量データとの回帰分析を行うことで収量予測モデルの構築も行った。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の大きさや、RGB の固定値で 2 値化を行っているため、天候による輝度変化や、落花も対象に含んでしまうなどロバスト性に問題があると考えられる。以上の背景を踏まえ、本研究では広範囲な空撮画像から土佐文旦の生理生態情報の一つである着花量を「見える化」する手法を提案する。具体的には、以下の 3 段階のアプローチをとる。第一に、物体認識モデルである YOLOv11 [12] を用いて、広域画像から解析対象となる樹体の位置を高速かつロバストに推定する。第二に、基盤モデルである SAM 2.1 (Segment Anything Model 2.1) [17] を導入し、検出された樹体の樹冠領域のみをセグメンテーションすることで、背景ノイズを除去する。第三に、抽出された樹冠画像に対して U-Net [15] を適用し、画素単位での着花領域の検出を行うことで、樹体ごとの着花量を定量化する。本手法により、従来のサンプリング調査に代わる、全数調査に基づいた高精度な収量予測および栽培管理支援システムの構築を目指す。

## 第 2 章

# 関連研究

本章では、IoP 関連研究、および本研究の基盤となるディープラーニング手法について記述する。

### 2.1 IoP (Internet of Plants) プロジェクトにおける生理生態情報の可視化

高知県では、施設園芸を対象として、作物の光合成、蒸散、栄養成長、生殖成長等の生理生態情報を「見える化」し、営農支援情報として「使える化」を行い、それらを産地全体で「共有化」する IoP (Internet of Plants) プロジェクトが推進されている [3, 6]。IoT 機器を用いて収集した環境情報や画像データを IoP クラウドに集約し、ハイブリッド AI モデル群を中核とした生理生態 AI エンジンによって解析することで、生産者に説明性の高い情報を提供している [7]。この研究は、ハウス内環境データと作物群落の定点画像を統合的に解析することで、群落光合成速度、蒸散速度、葉面積、群落体温といった生理指標に加え、栄養成長と生殖成長のバランスをリアルタイムで可視化することを可能にした。具体的な適用例として、果菜類であるナスにおいては、群落光合成情報に基づき、開花数、着果数、果実肥大の推移、および収穫予定日のリアルタイム推定が実現されている [7]。また、葉菜類であるニラにおいては、オートエンコーダ等の AI 技術を駆使し、単一の群落画像から草丈、生重、分けつ数を即時的に推定する手法が確立されている [7]。

## 2.2 露地果樹栽培における課題と画像解析手法

IoT プロジェクトにより施設園芸では高度な環境制御と生理生態情報の可視化が進む一方で、露地栽培においてはまだ十分な実装が進んでいないのが現状である。特に高知県の特産である露地土佐文旦等の果樹園は広大であり、個々の樹体の状態把握には多大な労力を要する。谷本ら [14] は、着花量や葉面積指数の簡易測定技術について検討しているが、広域かつ高頻度なデータ取得のためには、UAV 等で撮影された高解像度画像を自動解析する技術の確立が不可欠である。

画像解析による果実や樹体の検出には、リアルタイム性と検出精度に優れた YOLO (You Only Look Once) シリーズ [11] が広く用いられており、最新の YOLOv11[12] など進化を続けている。例えば、Koirala ら [9] は、YOLO を用いてマンゴーの果実検出を行い、高精度な収量推定が可能であることを示している。しかし、YOLO によるバウンディングボックス (BBOX) のみの検出では、背景の葉や枝が含まれるため、正確な樹冠面積や花密度の算出には限界がある。これに対し、画素レベルでの詳細な領域抽出には Semantic Segmentation 等のセグメンテーション手法が有効である。Dias ら [10] は、リンゴの花房検出に CNN を用いており、開花数と収量の相関分析においてセグメンテーションの有効性が示唆されている。また、医療画像分野で発展した U-Net[15] も、微細な構造の抽出に強力な性能を発揮する。さらに近年登場した Segment Anything Model (SAM) [16] やその改良版である SAM 2[17] は、プロンプトとして BBOX 等を与えることで、ゼロショットかつ高精度に任意の物体をセグメンテーション可能であり、農業画像のような複雑な背景を持つタスクへの応用が期待されている。既存研究の多くは、果実の直接検出または単一モデルによる解析に留まっている。これに対し本研究では、YOLOv11 による樹体検出、SAM による樹冠抽出、U-Net による花房抽出を階層的に組み合わせることで、UAV 空撮画像から土佐文旦の収量を予測する新たなモデルを提案する。

### 2.3 物体認識・セグメンテーション技術

#### 2.3.1 CNN (Convolutional Neural Networks)

CNN (畳み込みニューラルネットワーク) は、現在の画像認識技術の中核をなす深層学習モデルである。従来の画像処理手法では、2004年にLoweらにより提案されたSIFT [19] や2005年にDalal and Triggsらにより提案されたHOG[20]のようにエッジやテクスチャといった特徴量を人間が手動で設計する必要があったが、1998年にLeCunらにより提案されたCNNは畳み込み層 (Convolutional Layer) とプーリング層 (Pooling Layer) を多層に重ねることで、画像データから直接、高次元の特徴量を自動的に学習・抽出することが可能である [22]。特に、AlexNet[23] の登場以降、ILSVRCなどの画像認識コンペティションにおいて飛躍的な精度向上が達成され、現在では農業分野を含むあらゆる画像解析タスクにおいて、標準的な技術となっている。本研究で用いるYOLO[11] やU-Net[15]などのモデルも、基本構造としてこのCNNをバックボーンに採用しており、植物の複雑な形状や色情報を捉えるための基礎技術となっている。

#### 2.3.2 YOLO (You Only Look Once)

YOLO (You Only Look Once) は、Redmonら [11] によって提案されたリアルタイム物体検出アルゴリズムである。従来の2014年のR. Girshickら [13]R-CNNなどの手法が、関心領域の提案とクラス分類を別々の工程で行う2段階検出 (Two-stage detector) であったのに対し、YOLOはこれらを単一の回帰問題として解く1段階検出 (One-stage detector) を採用している。このアーキテクチャにより、画像全体から大域的な文脈情報を一度に考慮し、物体の位置矩形 (Bounding Box) とクラス確率をend-to-endで推定することが可能となり、極めて高速な処理を実現している。本研究では、広範囲な空撮画像から解析対象となるユズおよび土佐文旦の樹体位置を特定するための検出器として、YOLOの最新モデルであるYOLOv11 [12] を採用する。YOLOv11は初期のモデルと比較して、以下の点で本研究の目的に合致した学術的・技術的優位性を有する。第一に、アンカーフリー方式の採用で

## 2.3 物体認識・セグメンテーション技術

ある。従来手法では、検出対象の形状を想定したアンカーボックスの事前定義が必要であり、これが検出精度のボトルネックとなる場合があった。対して YOLOv11 は、特徴マップ上のグリッドセルから対象物の中心距離を直接予測する Center-based なアプローチをとる。これにより、UAV の撮影高度変化に伴う樹体のスケール変動や、剪定状態による形状のばらつきに対して、ハイパーパラメータの煩雑な調整なしに柔軟に対応できるロバスト性を獲得している。第二に、改良されたバックボーン構造と Attention 機構の導入である。detection and segmentation YOLOv11 では、特徴抽出層に新たな C3k2 ブロックを採用し、浅い層から深い層への情報の伝達効率を向上させている。加えて、C2PSA (Cross Stage Partial with Spatial Attention) モジュールを導入することで、画像内の重要な領域を強調し、背景のノイズを抑制する空間的な注意機構が強化されている。果樹園のような非構造的な環境において、複雑な背景から対象を正確に分離するためには、この高度な特徴表現能力が不可欠である。第三に、パラメータ効率の高さである。本研究で採用する YOLOv11n (Nano) は、分離型ヘッド構造により、クラス分類とバウンディングボックス回帰を独立して学習させることで、わずか 2.6M 程度のパラメータ数でありながら、旧来の大型モデルに匹敵する精度を達成している [12]。これは、本研究が目指す将来的なエッジデバイス実装やリアルタイム解析において極めて有利な特性である。以上の理由から、本研究では YOLOv11n を用いて樹体検出を行う。

### 2.3.3 SAM (Segment Anything Model)

SAM (Segment Anything Model) は、Kirillov ら [16] によって提案された、画像セグメンテーションのための基盤モデル (Foundation Model) である。SAM は 1,100 万枚の画像と 10 億以上のマスクからなる大規模データセット (SA-1B) で学習されており、追加の学習なしで未知の物体やドメインに対応できる強力な Zero-shot 性能を有する。モデルのアーキテクチャは、画像を処理する画像エンコーダ、プロンプト (点, ボックス, テキスト等) を処理するプロンプトエンコーダ, およびマスクエンコーダの 3 部構成となっている。一度画像エンコーダで画像を埋め込み表現に変換すれば、以降は軽量なプロンプトエンコーダと

## 2.3 物体認識・セグメンテーション技術

マスクエンコーダのみで推論が完結するため、異なるプロンプトに対してリアルタイムなマスク生成が可能である。本研究では、SAMの改良版であるSAM 2.1 [17]のLargeモデル(sam2.1l)を採用する。SAM 2.1における学術的に重要な改良点は、バックボーンネットワークにHiera (Hierarchical Vision Transformer) [18]を採用した点にある [17]。従来のViT (Vision Transformer) を用いたSAMと比較し、Hieraは階層的な特徴抽出を行うことで、画像の局所的詳細と大域的文脈を効率的に統合可能である。これにより、樹木の枝葉のような複雑かつ微細な境界を持つ対象に対しても、マスクの境界精度 (Boundary Quality) が大幅に向上している。また、検出器として軽量なYOLOv11nを採用した一方で、セグメンテーションにはパラメータ数の多いLargeモデルを選定した理由は、樹冠領域の抽出における特徴表現力を重視したためである。果樹園の環境下では、対象木と背景(草地や隣接木、防風林)が色彩的に類似しており、単純な閾値処理や低容量のモデルでは分離が困難である。Largeモデルの高いモデル容量 (Model Capacity) は、わずかなテクスチャの違いを捉え、枝葉によるオクルージョン (閉塞) を含む複雑な形状に対してもロバストな領域抽出を実現する。本提案手法では、前段のYOLOによって検出されたバウンディングボックスをプロンプトとしてSAM 2.1に入力することで、各樹体の位置特定と樹冠領域の分離を同時に達成する。

### 2.3.4 U-Net

U-Netは、Ronnebergerら [15]によって提案された、全層畳み込みネットワーク (Fully Convolutional Network: FCN) を拡張したセマンティックセグメンテーションモデルである。元来は、電子顕微鏡画像における細胞内構造の特定など、学習データが限定的でありながら、画素単位での高精細な認識が求められる医用画像処理のタスクのために開発されたモデルである。本モデルのアーキテクチャは、画像の文脈情報を捉えるための収縮パスと、正確な位置情報を復元するための拡張パスから構成され、全体として対称的なU字型の構造を成している。収縮パス (エンコーダ) では、畳み込み処理とプーリング層によるダウンサンプリングを繰り返すことで、画像の大域的な特徴、抽象的な情報を抽出する。しかし、こ

## 2.3 物体認識・セグメンテーション技術

の過程では受容野が広がる一方で空間解像度が低下するため、微細な物体の位置情報は失われる傾向にある。これを解決する U-Net 最大の特徴が、スキップ接続 (Skip Connection) である。これは、エンコーダ側の各解像度レベルで保持されている高解像度な特徴マップを、対応するデコーダ側の層へ直接入力し、チャンネル方向に結合する構成である。これにより、デコーダはアップサンプリングによって解像度を復元する際、深層で得られた意味的な情報と、浅層から転送された詳細な空間情報を同時に利用することが可能となる。本研究におけるカンキツの花は、画像全体に対して画素数が極めて少なく、かつ密集して存在する小物体である。一般的な CNN ではプーリング操作により消失しやすいこのような微細な特徴に対し、U-Net のスキップ接続は空間的な位置情報を保持し続けるため、有効に機能する。本手法では、YOLO, SAM によって抽出された樹冠領域セグメンテーション画像を U-Net に入力し、画素単位でのクラス分類を行うことで、花領域の定量的な計測を実現する。

## 第 3 章

# 提案手法

本章では、提案手法について示す。従来手法では、手動での樹冠領域の抽出を行っていたのに対して、本研究では、YOLO を持ちいた樹冠領域推定と SAM を用いた樹冠領域セグメンテーションを行うことで樹冠領域抽出の簡略化を行う。また従来手法では、抽出した領域内に RGB 閾値で画像を花と後景を分類したマスク画像を生成していた。これに対して提案手法では、U-Net を用いて花領域セグメンテーションを行うことで、ノイズに強い花量推定を行う。また、従来手法と同様に、推定花量データと実測収量データを持ちて予測モデルの構築するが、従来手法では線形モデルのみであったが非線形モデルを用いる。これらの手法を提案する。処理手順は以下の 4 段階で構成した。提案手法の全体図を図 3.1 に示す。

1. 樹冠領域の ROI (Region of Interest) : YOLOv11n による BBOX での樹冠認識
2. 樹冠領域のセグメンテーション : BBOX をプロンプトの入力として受け取り SAM による樹冠領域のセグメンテーションを行う。
3. 花領域の推定 : U-Net による花房のセマンティックを行い、推定花量を算出する。
4. 収量モデルの作成 : 1 樹体あたりの推定花量と 1 樹体あたりの果実の収量とで線形回帰、多項式回帰、サポートベクタ回帰で収量予測モデルを作成する。

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、対象樹木の花領域のみを解析することを可能し、1 樹体あたりの着花量を定量化できる。先行研究の手法で、『背景や撮影距離に影響されない着花量の定量化指標として、空撮画像中の樹冠投影面積に占める花の面積の割合を推定花量と定義した。』[谷本ら, 2025][14] とあるため、本研究でも、推定花量に同様の定義を用いる。

### 3.1 BBOX での樹冠領域推定

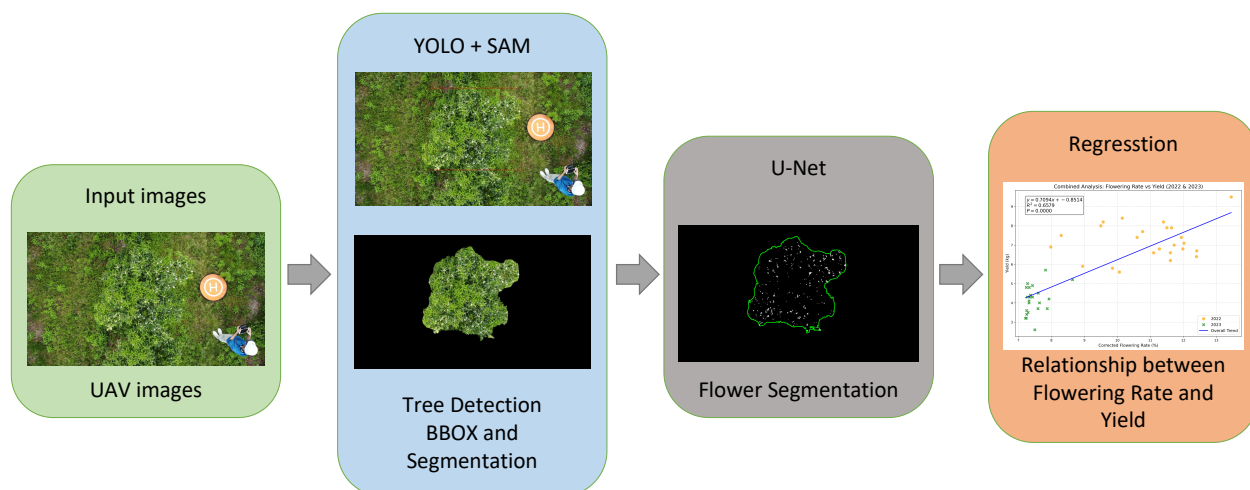


図 3.1: 提案手法の概要

### 3.1 BBOX での樹冠領域推定

検出モデルには YOLOv11n を選定した。本モデルを採用した理由は以下の 2 点である。第一に、YOLOv11 は従来の YOLOv8 等と比較してパラメータ効率が最適化されており、少ないパラメータ数で複雑なデータセット上でより高い mAP を達成している点である。第二に、モデルサイズである。本研究の樹冠領域推定モデルで使用するデータセット規模 (200 枚程度) において、過大なパラメータを持つモデル (Large 等) を使用することは過学習のリスクを高める。そのため、パラメータ数が約 2.6M と最小であり、学習データに対して適切なモデルの複雑さを持ち、かつ推論速度に優れる Nano モデルが最適であると判断した。また、今回対象とする樹体は各 UAV 空撮画像の中央に位置する樹体である。そこで検出された BBOX の内、中心座標が画像の中央 30 % の領域に BBOX のみをリストアップし、そのリストアップされた BBOX について、画像の中心点から BBOX の中心座標間のユークリッド距離が最も近いもの予測の BBOX としている。さらに SAM に入力として BBOX を入力するときは、元の BBOX のサイズ ( $w, h$ ) について 1.2 倍して受け渡す。これは、カンキツ樹体は形状が不規則であり、YOLO による矩形検出では、樹冠の周縁部にある微細な枝葉が BBOX の範囲外となる場合がある。SAM は入力された BBOX の範囲外をマスク候補として考慮しない特性があるため、BBOX が見切れていると正確な樹冠面積を算出でき

### 3.1 BBOX での樹冠領域推定

ない。そのため、YOLO の検出矩形に対して 1.2 倍のスケール係数を乗じ、樹体全体を確実に内包する領域 ROI を再定義した上でセグメンテーションを実行した。このマージン確保により、YOLO の検出誤差による欠損を防ぎ、SAM による精密な抽出を可能とした。

#### YOLO による処理過程

本手法において、YOLOv11n は単なる物体検出器としてだけでなく、後段のセグメンテーション (SAM) および花検出 (U-Net) が正確に動作するための初期解を提供する重要な役割を担う。その処理プロセスは、以下の 3 段階で構成される

入力画像  $I$  は、CNN バックボーンによって複数の解像度を持つ特徴マップ (Feature Pyramids) に変換される。YOLOv11 はアンカーフリーアーキテクチャを採用しているため、各特徴マップのグリッドセルにおいて、以下の 2 つの値を直接回帰によって予測する。

- **バウンディングボックス回帰**: グリッド中心からボックスの 4 辺 (上, 下, 左, 右) までの距離。
- **クラス分類スコア (confidence score)**: 検出対象が「樹木」である確率。

本研究では、推論設定において以下の学調整を行っている。高解像度推論 ( $1280 \times 1280$ )pixels: 標準的な  $640 \times 640$ pixels ではなく  $1280 \times 1280$ pixels での推論を行うことで、空撮画像における小物体 (Small Objects) への感度を向上させた。これにより、広域な UAV 空撮画像内の樹冠の端にある詳細な枝葉の特徴消失を防止している。低信頼度閾値 ( $conf = 0.1$ ): 信頼度閾値を 0.1 という低い値に設定することで、**再現率 (Recall)** を最大化する閾値を採用した。「見逃し (False Negative : FN)」を許容せず、過検出気味であっても候補をすべて列挙した上で、選別を後段の手順に委ねる設計である。

## 3.2 SAM による樹冠領域のセマンティックセグメンテーション

### Non-Maximum Suppression (NMS)

YOLO 等の検出モデルは、1 つの物体に対して多数の候補ボックスを出力する傾向がある。これらを整理するために **NMS (非極大値抑制)** が適用される。具体的には、IoU (Intersection over Union) に基づいて重複するボックスを統合し、最も信頼度の高いボックスのみを残すことで、1 本の樹木に対して 1 つのバウンディングボックスという一対一対応を確保する。

### 空間的ヒューリスティクスによる対象選定

空撮画像には隣接する樹木も写り込むが、本研究の解析対象は画像の「中心」に位置する樹体である。そこで、検出された候補群に対して**幾何学的な制約**を導入した。

全検出ボックスの集合  $\mathcal{B} = \{b_1, b_2, \dots, b_n\}$  の中から、画像中心座標  $(C_x, C_y)$  と各ボックスの中心座標  $(cx_i, cy_i)$  とのユークリッド距離  $d_i$  が最小となるボックス  $b_{\text{target}}$  を解析対象として選択する。

$$b_{\text{target}} = \operatorname{argmin}_{b_i \in \mathcal{B}} \sqrt{(cx_i - C_x)^2 + (cy_i - C_y)^2} \quad (3.1)$$

この処理により、画像の端に見切れている樹木などのノイズを排除し、データセット作成時の意図（中心の樹体を撮影対象とする）をアルゴリズム的に担保している。

## 3.2 SAM による樹冠領域のセマンティックセグメンテーション

YOLO によって検出されたバウンディングボックスは樹体の位置を大まかに示すものであるが、正確な開花率の算出には、BBOX 内に含まれる土壌や下草等の背景領域を除外し、樹冠の形状を画素レベルで正確に抽出する必要がある。そこで本研究では、Meta AI によって提案された基盤モデルである SAM の sam2.1l モデルを採用し、Coarse-to-Fine に基づ

## 3.2 SAMによる樹冠領域のセマンティックセグメンテーション

いたセグメンテーションを行う。本研究における SAM の処理手順は、以下の 3 段階の数理的モデルによって定式化される。

### 3.2.1 BBOX 位置情報プロンプトベースのマスク生成

SAM は、画像エンコーダ (Image Encoder)、プロンプトエンコーダ (Prompt Encoder)、およびマスクデコーダ (Mask Decoder) の 3 つの要素から構成される。入力画像  $I \in \mathbb{R}^{H \times W \times 3}$  は、まず画像エンコーダ (Hierarchical Vision Transformer) によって特徴量空間へ写像され、画像埋め込み  $Z_{\text{img}}$  が生成される。

一方、前段の YOLO で得られた BBOX :  $B_{\text{scaled}}$  は、プロンプトエンコーダに入力される。ここで、BBOX の座標情報は Positional Encoding によって高次元ベクトル  $Z_{\text{prompt}}$  へと変換される。これは、画像内の関心領域 ROI となり、空間的な制約として機能する。

マスクデコーダは、これら 2 つの埋め込み情報を受け取り、以下の写像関数  $\mathcal{F}_{\text{SAM}}$  によって樹冠領域の確率マップを生成する。

$$M_{\text{prob}} = \mathcal{F}_{\text{SAM}}(Z_{\text{img}}, Z_{\text{prompt}}) \quad (3.2)$$

ここで、本実装では推論オプションとして `retina_masks=True` を有効化している。これにより、デコーダの出力は低解像度の特徴マップ上ではなく、入力画像の解像度に合わせてアップサンプリングされ、樹冠周縁部の複雑な枝葉の形状を維持した高解像なバイナリマスク  $M_{\text{raw}}$  が生成される。

### 3.2.2 マスクの統合と精緻化

SAM は同一のプロンプトに対して曖昧な複数のマスク候補を出力する場合がある。本手法では、検出された全マスクの論理和をとることで初期マスクを形成する。しかし、この初期マスクには、枝葉の隙間に生じた微小なノイズや、樹冠から独立した誤検出領域が含まれる可能性がある。これを除去し、解析対象となる単一の樹冠領域を確定させるため、以下の画像処理アルゴリズム  $\mathcal{P}_{\text{refine}}$  を適用する。

### 3.3 U-Net による花領域 Semantic Segmentation

1. **モルフォロジー演算（オープニング）**：マスク画像に対してカーネルサイズ  $5 \times 5$  の構造要素を用いたオープニング処理（収縮後に膨張）を適用し，微小な孤立点やスパイクノイズを除去する．
2. **最大連結成分の抽出（LCC）**：ノイズ除去後のマスクから全ての輪郭を抽出し，その面積  $S$  が最大となる輪郭のみを保持する．これにより，主要な樹冠領域以外を背景として除去する．

最終的な樹冠マスク  $M_{\text{final}}$  は次式で表される．

$$M_{\text{final}} = \underset{c \in \mathcal{C}(\text{Open}(M_{\text{raw}}))}{\text{argmax}} S(c) \quad (3.3)$$

ここで， $\text{Open}(\cdot)$  はオープニング処理， $\mathcal{C}(\cdot)$  は輪郭抽出関数， $S(c)$  は輪郭  $c$  の面積を表す．この  $M_{\text{final}}$  を用いて画像  $I$  をクロッピング（Masking Cropping）することで，次段の U-Net による花検出処理への入力画像を生成する．実際に生成したセグメンテーション画像を図 3.2 に示す．

## 3.3 U-Net による花領域 Semantic Segmentation

SAM によって抽出された樹冠領域内部において，画素単位での花領域抽出を行うために，全層畳み込みニューラルネットワークである U-Net を採用した．本研究における U-Net の役割と推論手順は以下の通りである．

### 3.3.1 ネットワークアーキテクチャ

本研究で用いる U-Net は，画像内の文脈情報を抽出する収縮パスと，位置情報を復元する拡張パスからなる U 字型の Encoder-Decoder 構造を有する．今回構築した U-Net を図 3.3 に示す．

- **エンコーダ**：4 段階のダウンサンプリングを行い，画像の解像度を下げながら特徴量

### 3.3 U-Net による花領域 Semantic Segmentation



図 3.2: 樹冠領域 BBOX から生成したセグメンテーション画像

チャンネル数を最大 1024 まで拡張する。各ブロックは  $3 \times 3$  の畳み込み層、バッチ正規化 (Batch Normalization), および ReLU 活性化関数の繰り返しにより構成される。

- **デコーダ:** 転置畳み込み (Transposed Convolution) を用いて特徴マップをアップサンプリングする。この際、エンコーダ側の対応する解像度の特徴マップをスキップ接続 (Skip Connection) によって連結 (Concatenate : cat) することで、ダウンサンプリング過程で失われた空間的な詳細情報を補完する。

### 3.3 U-Net による花領域 Semantic Segmentation

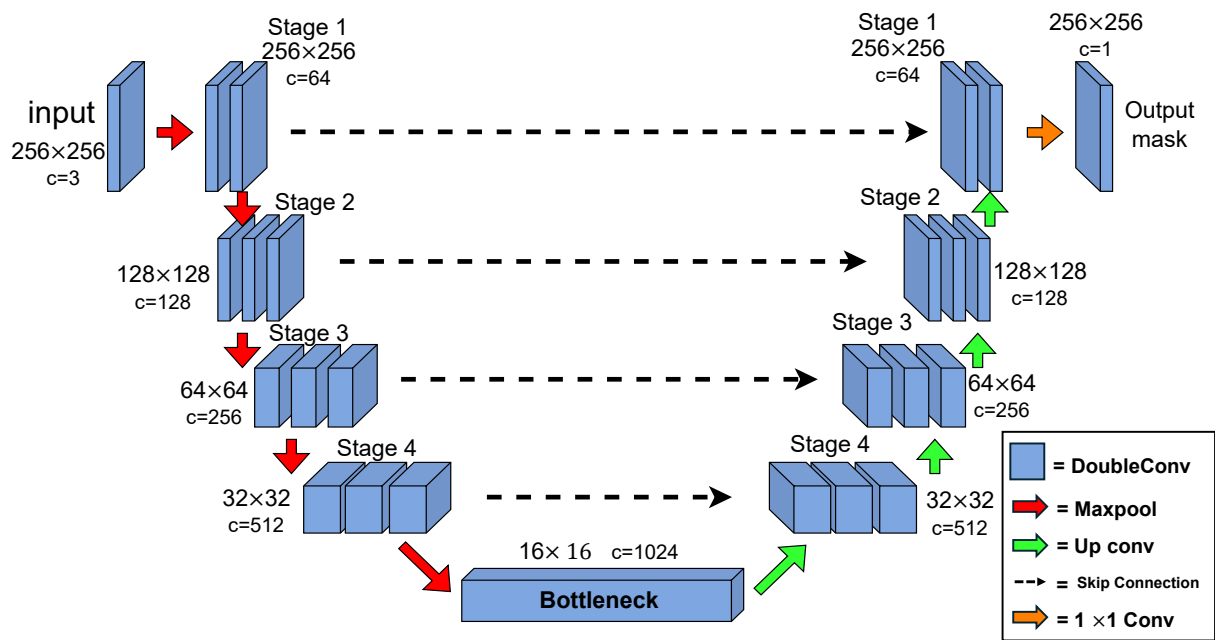


図 3.3: U-Net 構成図

最終層においては、 $1 \times 1$  の畳み込みによりチャンネル数をクラス数（本研究では花か否かの 2 値）に集約し、シグモイド関数により各画素が花である確率  $P_{flower} \in [0, 1]$  を出力する。

#### 3.3.2 損失関数の設計

画像内の着花領域は背景（葉や枝）と比較して面積が極めて小さいため、クラス不均衡（Class Imbalance）の問題が発生しやすい。そこで本研究では、画素ごとの分類誤差を評価する重み付きバイナリ交差エントロピー誤差（Weighted BCE Loss）と、領域の重なり具合を評価する Dice Loss を組み合わせた複合損失関数  $L_{total}$  を定義した。

$$L_{total} = L_{BCE}(y, \hat{y}) + L_{Dice}(y, \hat{y}) \quad (3.4)$$

ここで、 $L_{BCE}$  は正例（花）の重みを  $\lambda_{pos} = 10.0$  に設定することで検出漏れを抑制している。

$$L_{Dice} = 1 - \frac{2 \sum (y \cdot \hat{y}) + \epsilon}{\sum y + \sum \hat{y} + \epsilon} \quad (3.5)$$

$y$  は正解ラベル、 $\hat{y}$  は予測確率、 $\epsilon$  は数値安定化のための平滑化項である。

### 3.3 U-Net による花領域 Semantic Segmentation

#### 3.3.3 スライディングウィンドウによるパッチ推論

抽出された樹冠画像は U-Net の入力サイズ  $256 \times 256$  pixels より高解像度であり、かつサイズが可変であるため、そのまま U-Net に入力することは困難である。また、単純な画像のタイル分割では、パッチの境界付近でコンテキスト情報が分断され、セグメンテーション精度が低下する問題がある。

そこで本手法では、**オーバーラップを伴うスライディングウィンドウ推論**を実装した。具体的には、入力画像を固定サイズ ( $256 \times 256$  pixels) のパッチ  $I_{\text{patch}}$  に分割して推論を行うが、その際、隣接するパッチ間で領域の 50% (128 pixels) が重複するようにストライドを設定する。

画像上の座標  $(x, y)$  における最終的な予測確率  $P_{\text{map}}(x, y)$  は、その画素を含む全てのパッチからの予測値の平均として算出される。

$$P_{\text{map}}(x, y) = \frac{1}{N(x, y)} \sum_{k=1}^{N(x, y)} P^{(k)}(x, y) \quad (3.6)$$

ここで、 $P^{(k)}(x, y)$  は  $k$  番目のパッチによる予測確率、 $N(x, y)$  は座標  $(x, y)$  が推論された回数 (重複回数) を表す。この平滑化処理により、パッチ境界における不連続性を解消し、一貫性のあるセグメンテーション結果を得ている。

#### 3.3.4 最終的な花領域の決定

U-Net により得られた確率マップ  $P_{\text{map}}$  に対し、閾値処理および SAM マスクとの論理積演算を適用して、最終的な花領域マスク  $M_{\text{flower}}$  を生成する。

$$M_{\text{flower}} = (P_{\text{map}} > \tau) \wedge M_{\text{tree}} \quad (3.7)$$

ここで、 $\tau$  は信頼度閾値 (本実験では 0.8 に設定)、 $M_{\text{tree}}$  は SAM によって抽出された樹冠領域マスクである。背景領域を強制的に除外 (Masking) することで、樹冠外のノイズによる誤検出を防止し、樹冠内部の開花率算出の分母と分子の整合性を担保している。

### 3.3 U-Net による花領域 Semantic Segmentation

#### 3.3.5 着花率の定義と定量的評価

本提案手法における最終的な出力として、個々の樹体における推定花量として着花率を定量的に算出する。谷本らの先行研究 [14] により、背景や撮影距離に影響されない着花量の定量化指標として、空撮画像中の樹冠投影面積に占める花の面積の割合を推定花量と定義されているため本研究でも同様に、SAM によって抽出された樹冠領域の面積（画素数）を用いて花領域の面積を正規化することで、撮影状況の高度誤差による樹体の大きさの変動に影響されない密度指標として推定花量を定義した。

着花率  $R_{\text{flowering}}$  [%] は、以下の式により算出される。

$$R_{\text{flowering}} = \frac{N_{\text{flower}}}{N_{\text{tree}}} \times 100 \quad (3.8)$$

ここで、 $N_{\text{tree}}$  は SAM によって抽出された樹冠マスク  $M_{\text{tree}}$  内の総画素数、 $N_{\text{flower}}$  は樹冠領域内において U-Net が花として検出した画素数である。それぞれの画像内の画素面積は以下のように定義される。

$$N_{\text{tree}} = \sum_{x,y} M_{\text{tree}}(x,y) \quad (3.9)$$

$$N_{\text{flower}} = \sum_{x,y} (M_{\text{flower}}(x,y) \wedge M_{\text{tree}}(x,y)) \quad (3.10)$$

ここで  $\wedge$  は論理積を表す。上式に示すように、花領域の算出において SAM の樹冠マスクとの論理積をとることで、樹冠外（地面や他の樹体、背景）に存在する白色のノイズ（落花、石等）が計算に含まれることを防ぎ、純粹に樹上の花のみを評価対象とする。上記の手順で出力した結果の画像を図 3.4 に示す。ピンク色でマスクされた領域がモデルの予測領域である。また、推定花量の可視化画像を図 3.5 に示す。緑色で囲まれた領域が YOLO + SAM で出力したセグメンテーション結果である。その領域内に存在する白い領域が U-Net で出力した花量である。

### 3.4 推定花量に基づく収量予測モデルの構築

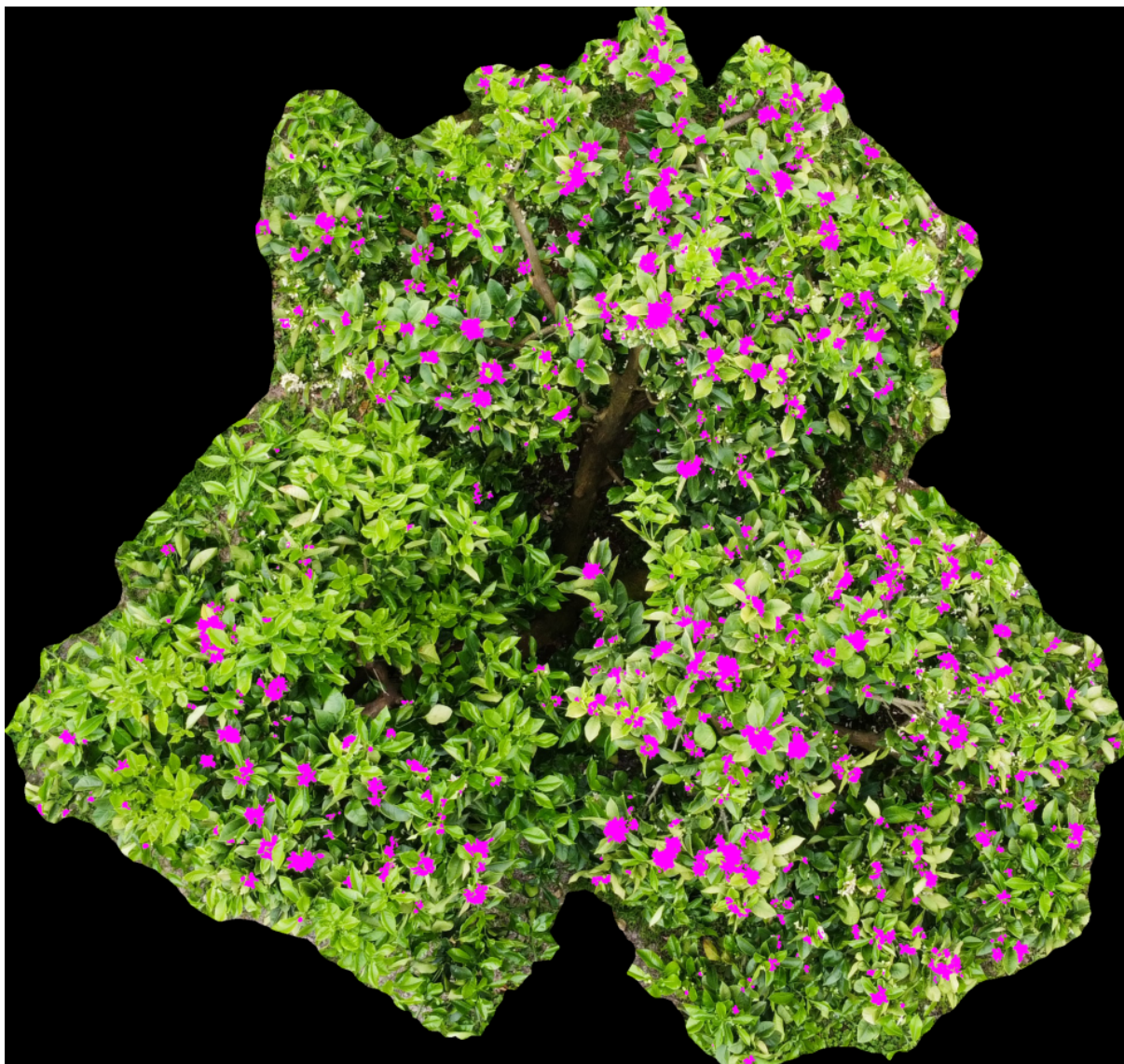


図 3.4: 花領域セグメンテーション結果画像

### 3.4 推定花量に基づく収量予測モデルの構築

算出された推定花量から最終的な文旦の収量 ( $kg/m^3$ ) を推定するために、提案手法では推定花量  $x$  を入力、収量  $y$  を出力とする回帰モデルを構築する。植物生理学的な観点およびデータ分布の特性を考慮し、以下の 3 種類のモデルを検討対象とした。

### 3.4 推定花量に基づく収量予測モデルの構築

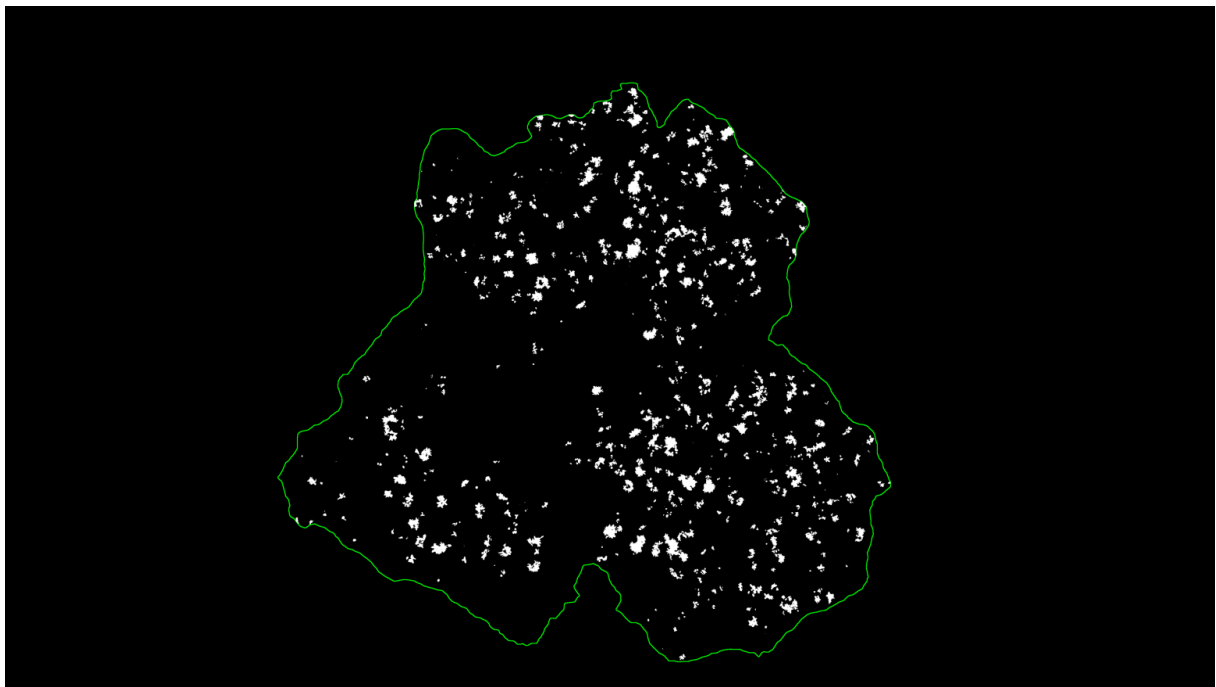


図 3.5: 推定花量の可視化画像

#### 3.4.1 線形回帰モデル (Linear Regression)

着花量と収量が単純な比例関係にあると仮定したベースラインモデルである。最小二乗法により、以下の関数のパラメータ  $w_0, w_1$  を決定する。

$$y = w_0 + w_1x \quad (3.11)$$

このモデルは、着花数の増加が常に収量増加に寄与する場合に有効である。

#### 3.4.2 2次多項式回帰モデル (Polynomial Regression)

果樹栽培においては、着花数が多すぎると生理落下が発生したり、果実一個あたりの肥大が抑制されたりする着果負担の影響が生じる可能性がある。このような飽和特性やピーク特性を表現するために、2次の項を含むモデルを採用した。

$$y = w_0 + w_1x + w_2x^2 \quad (3.12)$$

ここで  $w_2 < 0$  と推定された場合、過剰な着花が収量に対して負の影響、または増加の鈍化を与えることを示唆する。

## 3.4 推定花量に基づく収量予測モデルの構築

### 3.4.3 サポートベクター回帰 (Support Vector Regression)

露地栽培における外的要因を考慮するため、より複雑な非線形関係や、外れ値に対するロバスト性を考慮し、サポートベクター回帰 (SVR) を採用した。カーネル関数には動径基底関数を用い、以下の最適化問題を解くことでモデルを学習する。

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (3.13)$$

本手法では、特徴量の標準化を行った上で、正則化パラメータ  $C$  およびカーネル係数  $\gamma$  をグリッドサーチにより調整する。

## 第 4 章

# 実験内容

本章で研究で行った実験について示す。

### 4.1 樹冠領域推定モデルの評価実験

構築した樹冠領域推定モデル (YOLO) の汎化性能を評価する実験について以下に記述する。本研究で使用する UAV 空撮画像データセットは、高知県農業技術センター果樹試験所露地圃場 (以下、場内) に間隔  $4.0\text{m} \times 4.0\text{m}$  で植えられた土佐分担を供試した。樹体の撮影は、2021 年 4 月 27 日 (開花 1), 5 月 4 日 (開花 4~6 分, 7 日 (満開 (開花 8 分) ~落弁 1 分), 2022 年 5 月 1 日 (満開), 6 日 (開花終期 (落弁 5 分)) および 2023 年 4 月 26 日 (開花 1 分), 29 日 (開花 2 分), 5 月 4 日 (開花 6~7 分), 5 日 (満開) のいずれも曇天日の 6:30~17:00 の無風時に行った。また、場内に同様に植えられたユズの樹体も同様に撮影を行った。この開花割合は場内全体を達観で調査し高知県農業技術センターの職員が判断したものである。空撮用 UAV の Mavic mini (DJI 製) を用いて地上高 7.5~9.0m から、空撮影時表面が撮影地表面が横 13.3~15.9 m, 縦 10.0~11.9 m の範囲で、樹体に対して鉛直方向に撮影された画像である。撮影高度の差が 1.5m 程度あるが推定花量の定義を画像内の樹冠投影面積を算出し、面積内の花領域の占める割合で定義しているため、影響は小さいと判断した。画像の解像度は  $4000 \times 2250\text{pixels}$  である。本研究で対象としている着花量は土佐分担の花房を対象としているが、樹冠領域推定モデルのデータセットには、土佐分担と同じく高知県の特産であるユズ樹体空撮画像も含まれる。上記の内容で撮影された画像全 233 枚を訓練データ 173 枚、検証データ 53 枚に分割する。訓練データで YOLOv11n.pt

## 4.1 樹冠領域推定モデルの評価実験

に以下の条件で学習を行う。実際に使用したデータセットの例を図 4.1 に示す。



図 4.1: 樹冠領域推定モデルの学習データセット

### 4.1.1 モデル学習条件とハイパーパラメータ

初期重みには大規模データセット (COCO dataset) で事前学習されたパラメータを使用し、本研究で構築した樹体データセットの train データに対して転移学習を行った。使用する GPU は NVIDIA RTX 6000 Ada Generation である。(これ以降の実験も同様の GPU を使用する) 最適化アルゴリズムには、自動的に決定した確率的勾配降下法 (SGD) を採用した。過学習を抑制し汎化性能を高めるため、Weight Decay (重み減衰) および Mosaic データ拡張等を適用している。詳細なハイパーパラメータ設定を表 4.1 に示す。

学習データの多様性を確保し、屋外環境における照明変化や樹形のばらつきに対してロバストなモデルを構築するため、学習時に以下のデータ拡張を適用した。

- 幾何学的変換: 左右反転 (確率 0.5), スケーリング ( $\pm 50\%$ ), 平行移動 ( $\pm 10\%$ )
- 測光的変換: HSV 色空間における色相 ( $\pm 1.5\%$ ), 彩度 ( $\pm 70\%$ ), 明度 ( $\pm 40\%$ ) の

## 4.1 樹冠領域推定モデルの評価実験

表 4.1: Hyperparameters used for training YOLO11n

Parameter	Value
Model Architecture	YOLO11n (Nano)
Pretrained Weights	COCO
Input Resolution	1280 × 1280 pixels
Epochs	100
Batch Size	4
Optimizer	Auto(SGD)
Initial Learning Rate ( $lr_0$ )	0.01
Final Learning Rate ( $lr_f$ )	0.01
Momentum	0.937
Weight Decay	0.0005
Warmup Epochs	3.0

ランダム変化

- **Mosaic Augmentation:** 4 枚の画像をランダムに組み合わせて 1 枚の学習画像を生成する手法（確率は 1.0, ただし学習終了の 10 エポック前で無効化）

構築した樹形認識モデル（YOLO）の汎化性能を検証するために、学習時に使用しなかった検証データセットを用いた定量評価を行った。

### 4.1.2 推論設定

学習において検証損失が最小となった最適な重みパラメータ (best.pt) をロードし、モデルの初期化を行った。ドローン空撮画像における樹冠の詳細な特徴を捉えるため、入力画像の解像度は学習時と同様に 1280 × 1280 ピクセルに設定した。また、検証プロセスにおける信頼度閾値や NMS の IoU 閾値については、YOLO の標準的な検証プロトコルに従い、適

## 4.2 花領域検出モデルの評価実験

合率-再現率曲線 (PR Curve) を算出するために可変的な閾値を用いて評価した。

### 4.1.3 評価指標

検出精度の定量的評価には、YOLO の予測値は、真陽性 (TP)、偽陽性 (FP)、真陰性 (TN)、偽陰性 (FN) の 4 種類の評価基準で出力される、評価基準を使用する。関心領域 ROI の推定する樹体認識モデルの性能比較には、以下の Precision, Recall, AP@50 を用いる。

- **適合率 (Precision,  $P$ ):** モデルが樹木として検出した領域のうち、正しく樹木であった割合。

$$P = \frac{TP}{TP + FP} \quad (4.1)$$

ここで、 $TP$  (True Positive) は正検出数、 $FP$  (False Positive) は誤検出数を表す。

- **再現率 (Recall,  $R$ ):** 本来検出すべき樹木のうち、モデルが正しく検出できた割合。

$$R = \frac{TP}{TP + FN} \quad (4.2)$$

ここで、 $FN$  (False Negative) は未検出数を表す。

- **AP@50 (Average Precision@50):** 適合率と再現率の調和平均に基づく総合的な評価指標である。予測領域の BBOX と正解領域の BBOX の IoU (Intersection over Union) が 0.5 以上と場合を正解とした際の平均適合率である。

これらの指標を用いることで、樹冠領域の推論の正確性を包括的に評価する。

## 4.2 花領域検出モデルの評価実験

### 4.2.1 データセットの構築

本実験では、UAV 空撮画像から樹冠中心部を高解像度で切り出し、さらに小領域 (パッチ) に分割した画像データセットを構築した。元画像 (JPEG 形式, 解像度  $4000 \times 2250$  pixels, 72 dpi) からの具体的な生成手順は以下の通りである。

## 4.2 花領域検出モデルの評価実験

1. **樹木個体の選定:** 各撮影日において、解析対象とする樹木（10 個体または 12 個体）の画像を無作為に選出した。
2. **中心領域（親画像）の抽出:** 選出された各画像に対し、樹体中心部を含む  $1200 \times 1200$  pixels の正方形領域を切り出した（以下、親画像とする）。
3. **パッチ画像の生成:** 各親画像を縦横均等に 8 分割（グリッド分割： $8 \times 8 = 64$ ）し、サイズ  $150 \times 150$  pixels の小領域画像（以下、パッチ画像）を生成した。
4. **データセットのサンプリング:** 1 枚の親画像から生成された 64 枚のパッチ画像のうち、無作為に 10 枚を抽出した。これにより、1 撮影日あたり合計 100 枚または 120 枚のパッチ画像からなるデータセットを構築した。
5. **正解マスク画像の作成:** 生成した各パッチ画像に対して、手作業で花領域をアノテーション（白色描画）した正解マスク画像を作成した。

構築したデータセットは、パッチ画像と対応する正解マスク画像を 1 組として、学習用 (Train) と評価用 (Test) に 8 : 2 の割合でランダムに分割した。最終的なデータ数は、学習データ 943 枚、テストデータ 236 枚の合計 1179 枚である。本実験では、学習データを用いて U-Net モデルの学習を行い、テストデータを用いて提案手法および比較手法の性能評価を行った。

### 4.2.2 比較・評価手法

着花領域の検出性能を多角的に評価するため、提案手法 (U-Net アンサンブル) に加え、従来手法である色情報に基づく 3 種類の手法と比較を行った。

#### 提案手法: U-Net アンサンブル

5-fold 交差検証によって学習された 5 つの U-Net モデルを用いたアンサンブル推論を行う。各モデルが出力した確率マップの平均値を算出し、閾値  $T_{unet} = 0.8$  を適用して最終的な着花マスク  $M_{unet}$  を以下式 (4.3) で生成した。

## 4.2 花領域検出モデルの評価実験

$$M_{unet}(x, y) = \begin{cases} 1 & \text{if } \frac{1}{N} \sum_{i=1}^N P_i(x, y) > 0.8 \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

ここで、 $N = 5$  はモデル数、 $P_i(x, y)$  は  $i$  番目のモデルによる推論確率を表す。

### 比較手法 1: NTSC 加重グレースケール法適用後, RGB 色閾値 2 値化処理

NTSC 係数を用いて画像を以下の式 (4.4) でグレースケール化し, 以下の式 (4.5) で閾値処理を行う手法である。画像サイズはデータセット同様  $150 \times 150$  pixels で評価した。

$$Y = 0.2989R + 0.5866G + 0.1145B \quad (4.4)$$

$$M_{ntsc}(x, y) = \begin{cases} 1 & \text{if } Y(x, y) > \frac{195}{255} \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

### 比較手法 2: RGB 閾値法

色空間変換を行わず, RGB 各チャネルの画素値が閾値を超えている領域を花として検出する単純な手法である。使用する閾値は RGB 各チャネルが 195 を超えている場合 (4.6), Blue チャネルの値が 155 を超えている場合 (4.7) である。この 2 つの閾値は従来手法で性能が良かった 2 つの閾値である。

$$M_{rgb}(x, y) = \begin{cases} 1 & \text{if } (R > 195) \wedge (G > 195) \wedge (B > 195) \text{flower} \\ 0 & \text{Back ground} \end{cases} \quad (4.6)$$

$$M_b(x, y) = \begin{cases} 1 & \text{if } (B > 155) \text{flower} \\ 0 & \text{Back ground} \end{cases} \quad (4.7)$$

### 4.2.3 評価指標

各手法の性能を定量化するため, 以下の 2 つの指標を採用した。

## 4.2 花領域検出モデルの評価実験

### Intersection over Union (IoU)

IoU(Intersection over Union)とは、物体検出を行う際に、利用される正解の領域と予測した領域の重なりを定量化する指標である。IoUの算出方法を以下の図4.2に示す。IoUは、正解領域と予測領域の領域の共通部分を2つの領域の和集合で割ったものである。本実験の正解領域とは、アノテーションで作成した花領域を手作業で塗りつぶしたマスク画像である。予測領域は、モデルが花と予測した領域である。IoUは0から1の値をとり、1に近づくほど正解の領域と予測した領域の重なりが大きいことを示す。本実験でのIoUの閾値は0.8に設定し、この閾値を超えた予測値のみを陽性に分類する。データセット全体での平均

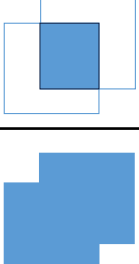
$$IOU = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$


図 4.2: IoU の計算式

値としてIoUを算出し、形状認識の正確さを評価する。

### 面積相関解析 (Regression Analysis)

検出された着花面積(ピクセル数)が、正解データの面積とどの程度相関しているかを検証するために線形回帰分析を行った。画像  $k$  における予測面積  $S_{\text{pred}}^{(k)}$  と正解面積  $S_{\text{gt}}^{(k)}$  の関係を以下の線形モデルで近似し、自由度調整済み決定係数 ( $R^2$ ) および赤池情報量基準 (AIC) を用いてモデルの当てはまりの良さを評価した。

$$S_{\text{gt}} = \alpha \cdot S_{\text{pred}} + \beta \quad (4.8)$$

$R^2$  値が1に近いほど、その手法による面積推定値が信頼できることを示唆する。

## 4.3 推定花量に基づく収量予測モデルの評価実験

### 4.2.4 実験条件の実装詳細

U-Net 用には  $256 \times 256$  ピクセル、色閾値 2 値化処理手法用には  $150 \times 150$  ピクセルへと動的にリサイズ処理を行った。

## 4.3 推定花量に基づく収量予測モデルの評価実験

前節までに算出された推定花量を用いて、実際の土佐分担の収量を予測する回帰モデルの構築実験を行った。本実験では、従来手法である RGB 閾値法と、提案手法である U-Net 法のそれぞれから得られた推定花量を使用し、収量との相関関係および回帰モデルの適合度を比較検証した。

### 4.3.1 データセットの前処理と補正

解析には、2022 年および 2023 年の 2 カ年分のデータを使用した。実際に使用する収量データを表 4.2 に示す。

画像解析によって得られた推定花量  $x_{raw}$  は、一次補正を行う。補正後の補正推定花量  $x_{calib}$  は以下の式 (4.9) で定義される。

$$x_{calib} = \alpha \cdot x_{raw} + \beta \quad (4.9)$$

ここで、 $\alpha$  は傾き、 $\beta$  は切片を表す。各手法における補正係数は、前節の実験結果を線形回帰で算出した補正式に基づき表 4.3 の通り設定した。また、 $3\sigma$  法により、極端な外れ値である個体 (C-11, C-27) は外れ値として解析から除外した。

### 4.3.2 回帰モデルの選定

推定花量と収量の関係性をモデル化するため、以下の 3 種類の回帰アルゴリズムを用いて適合度を比較した。

### 4.3 推定花量に基づく収量予測モデルの評価実験

表 4.2: 2022 年および 2023 年の収量データ (単位: kg/m<sup>3</sup>)

ID	2022	2023	ID	2022	2023
C-1	6.8	5.7	C-15	7.7	4.4
C-2	5.9	4.0	C-16	7.4	3.2
C-4	5.6	4.9	C-17	7.1	4.3
C-5	6.2	4.3	C-18	6.6	2.6
C-6	6.4	4.8	C-19	6.9	5.2
C-7	8.0	3.7	C-20	8.4	4.3
C-8	8.2	4.5	C-21	6.7	3.6
C-9	6.6	4.0	C-22	9.5	3.2
C-10	7.0	4.8	C-23	6.8	3.5
C-11	6.8	1.9	C-24	7.9	3.4
C-12	7.9	4.1	C-25	8.2	5.0
C-13	7.5	4.2	C-26	5.8	4.3
C-14	7.4	3.7	C-27	9.9	5.0

表 4.3: 着花率の補正係数

Method	Slope ( $\alpha$ )	Intercept ( $\beta$ )
RGB Threshold	0.538	-133.6
U-Net (Proposed)	0.878	7.2

1. **線形回帰:** 着花量と収量が単純な比例関係にあると仮定し, 最小二乗法によりパラメータを推定した.

$$y = w_1x + w_0 \quad (4.10)$$

2. **二次多項式回帰:** 着花過多による着果負担や摘果の影響 (収量の飽和傾向) を考慮する

#### 4.3 推定花量に基づく収量予測モデルの評価実験

ため、二次関数によるフィッティングを行った。

$$y = w_2x^2 + w_1x + w_0 \quad (4.11)$$

3. **Support Vector Regression (SVR)**: より複雑な非線形関係を捉えるため、動径基底関数 (RBF) カーネルを用いた SVR を採用した。ハイパーパラメータは、グリッドサーチに基づき  $C = 100$ ,  $\gamma = 0.1$ ,  $\epsilon = 1.0$  に設定した。

##### 4.3.3 評価指標

各モデルの性能評価には、決定係数 (Coefficient of Determination,  $R^2$ ) を用いた。

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (4.12)$$

ここで、 $y_i$  は実測収量、 $\hat{y}_i$  はモデルによる予測収量、 $\bar{y}$  は実測収量の平均値を表す。 $R^2$  が 1 に近いほど、算出された着花率が収量の変動を正しく説明できていることを示す。本実験では、RGB 法と U-Net 法それぞれにおいて、どの回帰モデルが最も高い  $R^2$  を示すかを検証した。

# 第 5 章

## 実験結果と考察

### 5.1 樹冠領域推定モデルの評価実験結果

学習済み樹冠領域推定モデル (best.pt) を用いて、検証データセット 53 枚に対する評価を行った結果を表 5.1 に示す。実験の結果、適合率 (Precision) は 0.997, 再現率 (Recall) は 1.000, AP@50 は 0.995 となり、場内における樹冠領域推定において十分な精度が得られた。

表 5.1: 樹冠領域推定モデル (best.pt) の評価結果

Precision	Recall	mAP@0.5
0.997	1.000	0.995

図 5.1 に、本モデルによる実際の推論結果の一例を示す。図中の赤色のバウンディングボックスはモデルが検出した樹冠領域を示している。

結果画像を確認すると、隣接する樹木同士の間隔が 4.0m で植栽されている環境では、個々の樹冠を分離して認識できていることがわかる。また、撮影条件である曇天日では、環境影響を受けずに樹冠領域を正しく推論できている。

#### 5.1.1 考察

本実験では、学習データが 173 枚と比較的少数であったが、Mosaic データ拡張や HSV 色空間での幾何学的・測光的変換を適用したことで、過学習を抑制し高い汎化性能が得られ

## 5.2 花領域検出モデルの評価実験結果



図 5.1: 樹冠領域推定モデル (YOLOv11nano) の出力画像

たとえられる。特に Mosaic 拡張は、様々なスケールの樹冠特徴を学習することに寄与し、大小異なる樹形の認識に効果を発揮したと推察される。モデルアーキテクチャとして軽量な YOLO11n を採用し、COCO データセットによる事前学習重みを用いた転移学習を行ったことも、早期の収束と精度の向上に寄与した。

## 5.2 花領域検出モデルの評価実験結果

テストデータ 236 枚を用いた各手法の IoU の比較結果を表 5.2 に示す。提案手法である U-Net アンサンブルは、比較手法の中で最も高い IoU 値 (0.722) を示した。また、決定係数でも比較手法の中で最も良い 0.985 を示した。AIC においても 3436 であり最もバランスの良い回帰モデルあることを示した。この結果により、提案手法である U-Net を用いた花領域セグメンテーションは葉や日光の照り返し、落花などの背景ノイズに影響を受けずに花領域の検出を可能であることが示された。

各手法により算出された着花予測面積と、正解マスクの面積との相関関係を図 5.2 に示す。

## 5.2 花領域検出モデルの評価実験結果

表 5.2: Comparison of segmentation performance (IoU)

Method	IoU	$R^2$	AIC
Proposed (U-Net Ensemble)	<b>0.722</b>	<b>0.985</b>	<b>3436</b>
Comp. 1 (NTSC + RGB195Threshold)	0.262	0.514	3753
Comp. 2 (RGB195 Threshold)	0.329	0.794	3551
Comp. 3 (Blue155 Threshold)	0.448	0.857	3464

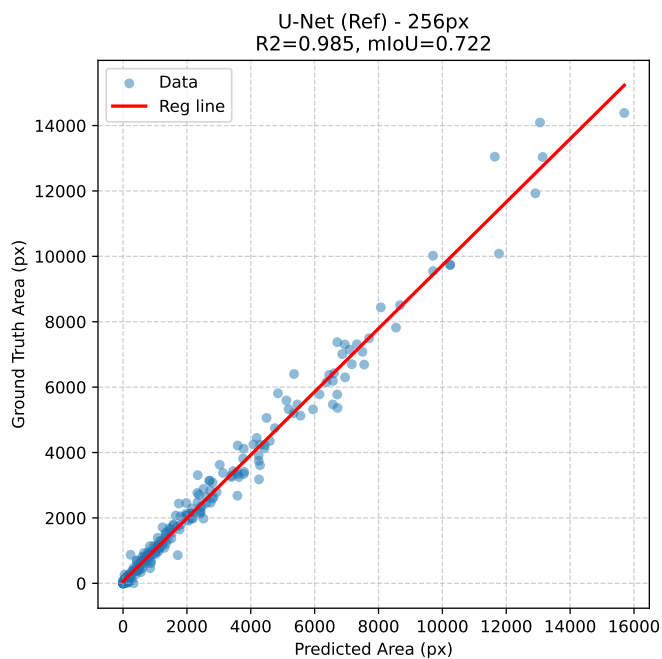
図 5.2 の散布図を確認すると、比較手法（特に (c) および (d)）は、回帰直線よりも上側にプロットが分布する傾向が見られた。これは、実際の花面積よりも予測面積を大きく算出していることを意味しており、葉の光沢や地面の雑草などを花として過検出していることが示唆される。一方で、提案手法 (a) はデータポイントが回帰直線付近に密集しており、過検出および検出漏れが少なく、個体ごとの着花量の変動を正確に捉えていることが確認できる。

図 5.3 に、各手法による検出結果マスク画像の視覚的な比較を示す。可視化の図より、提案手法はノイズを含まず検出できていることが確認できる。また従来手法は色の明るい葉の誤認識や暗い花の見逃しが確認された。

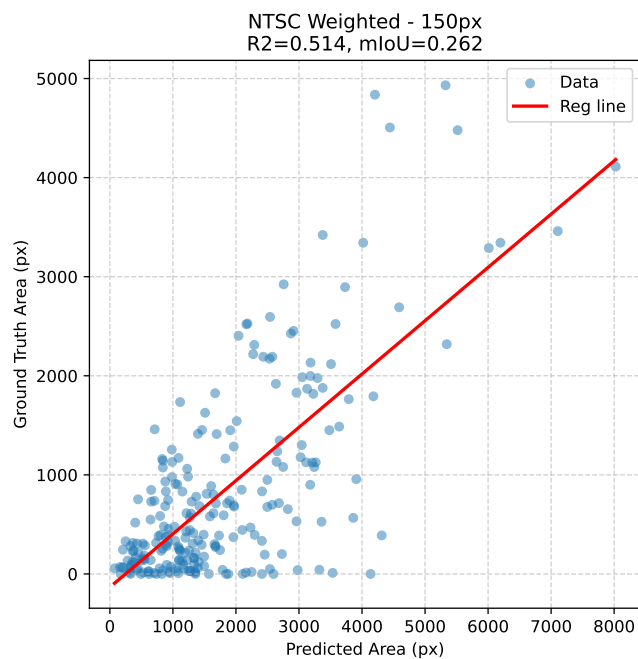
### 5.2.1 考察

結果より、比較手法 1 (NTSC) および比較手法 2 (RGB 閾値) の IoU が 0.262, 0.329 と低迷した要因として、露地環境特有の照明条件の変動が挙げられる。特に、晴天時における葉への直射日光の反射（ハレーション）は、RGB すべての画素値を上昇させるため、色閾値処理では白い花との判別が困難であったと考えられる。また、比較手法 3 (Blue チャンネル閾値) は、植物の葉が青色光を吸収する特性を持つため、葉と花のコントラストがつきやすく、比較手法の中では最も良好な結果 (IoU 0.448) を示した。しかし、それでも提案手法の IoU 0.722 には遠く及ばなかった。これに対し、提案手法である U-Net アンサンブルは、局所的な輝度値だけでなく、畳み込み処理によって対象物の形状やテクスチャといった

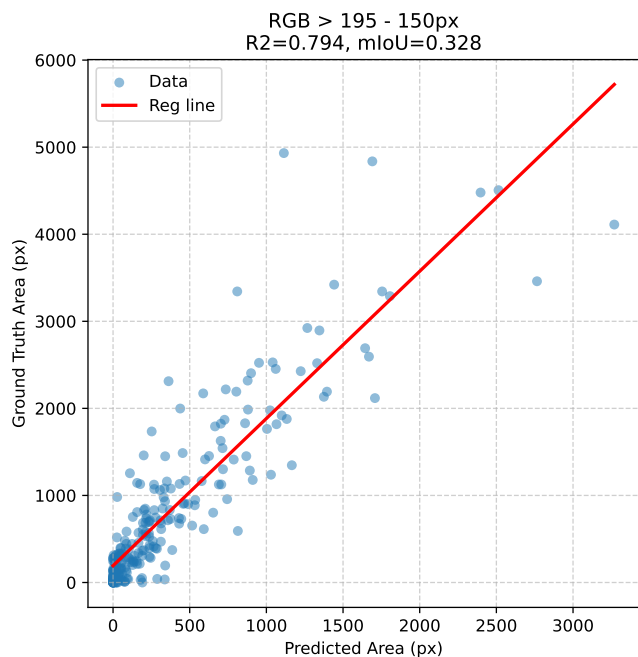
## 5.2 花領域検出モデルの評価実験結果



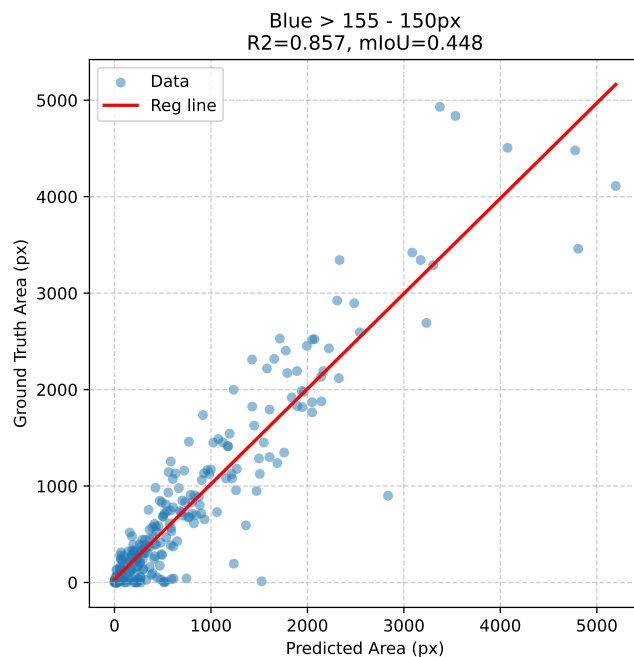
(a) U-Net



(b) NTSC + RGB195



(c) RGB195



(d) Blue155

図 5.2: U-Net(a) と従来手法 (b)-(d) の散布図

## 5.2 花領域検出モデルの評価実験結果

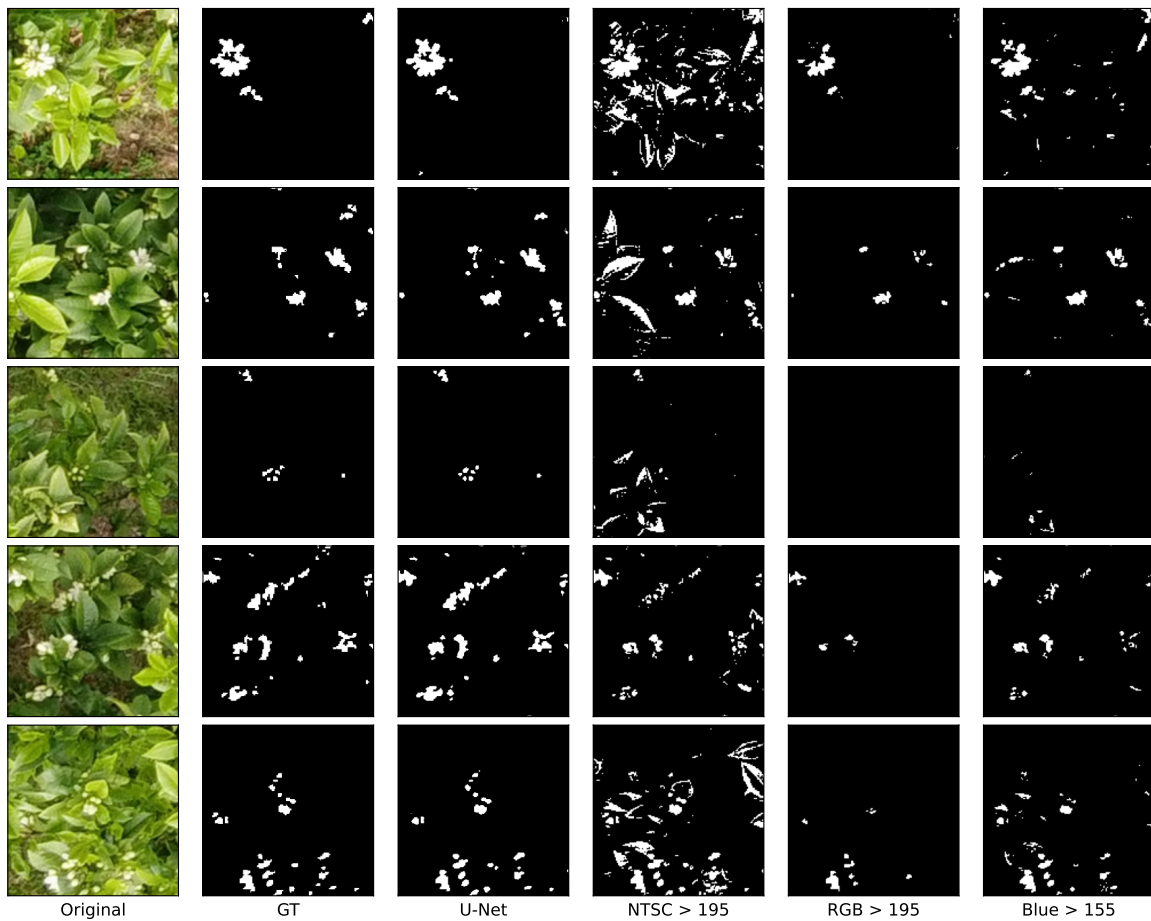


図 5.3: 提案手法と従来手法の比較結果

空間的な文脈情報を学習している。その結果、色情報だけでは区別がつかない強く反射した葉と花卉の違いを形状特徴から識別でき、背景ノイズを効果的に抑制できたと考えられる。表 5.2 において、比較手法 3 は決定係数 ( $R^2$ ) が 0.857 と比較的高い値を示したものの、IoU は 0.448 に留まった。これは、面積ベースの回帰分析だけでは検出性能を正しく評価できない危険性を示している。例えば、背景ノイズを大量に誤検出しつつ、実際の小さな蕾を検出漏れした場合でも、総画素数が偶然正解データと近ければ  $R^2$  は高くなる。しかし、このようなモデルは着花分布（花の密集度合い）の解析には利用できない。IoU と  $R^2$  の双方で高い値を示した提案手法のみが、着花位置と総量の両方を正しく定量化できており、栽培管理への応用において信頼性が高いと言える。

## 5.3 推定花量に基づく収量予測モデルの評価実験結果

### 5.3.1 収量予測精度の定量的評価

RGB 閾値法および提案手法 (U-Net 法) によって算出された補正推定花量を用い, 3 種類の回帰モデル (線形回帰, 二次多項式回帰, SVR) を構築した際の自由度調整済み決定係数  $R^2$  の比較を表 5.3 に示す.

表より, 二次多項式回帰を除く回帰モデルにおいて, RGB 閾値法と比較して提案手法である U-Net 法が高い  $R^2$  を示した. しかし, RGB 法と U-Net 法  $R^2$  の差は 0.01 程度であった. 回帰モデル間での比較では, 両手法ともに線形回帰よりも非線形モデル (二次多項式, SVR) の適合度が高く, 特に SVR が最も良好な結果を示した.

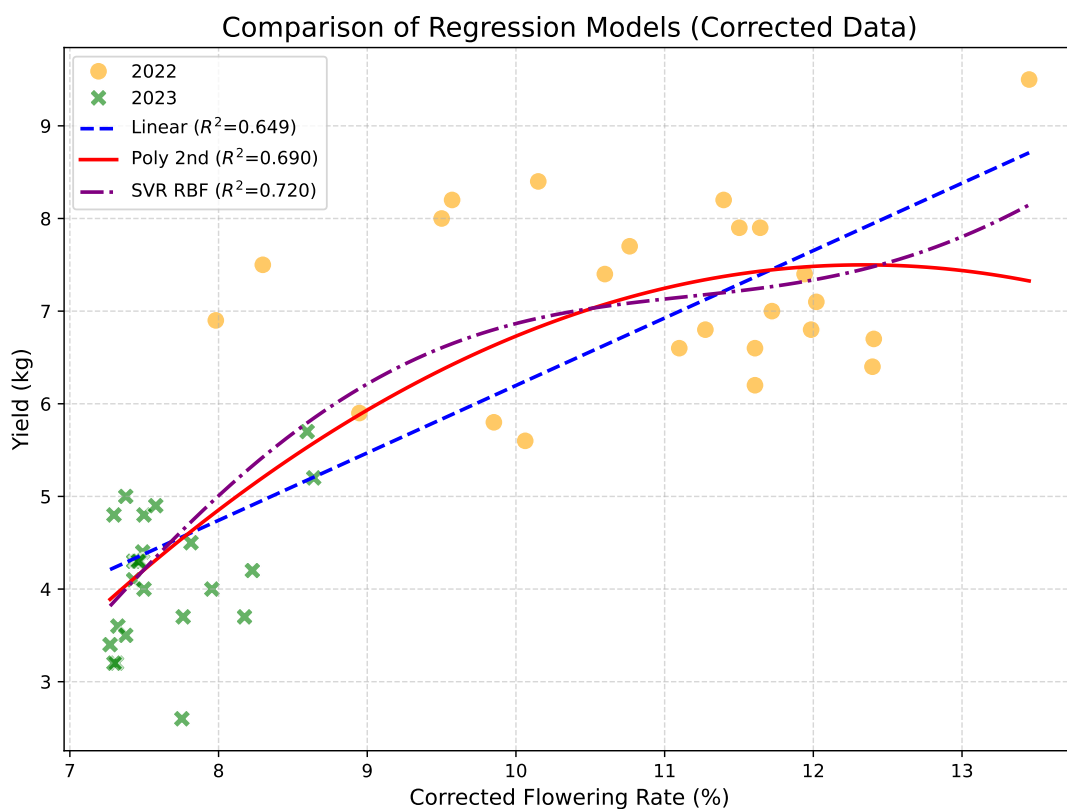
表 5.3: 各手法および回帰モデルにおける自由度調整済み決定係数 ( $R^2$ ) の比較

Estimation Method	Linear	Polynomial (2nd)	SVR
RGB Threshold	0.623	<b>0.679</b>	0.707
<b>U-Net (Proposed)</b>	<b>0.641</b>	0.676	<b>0.714</b>

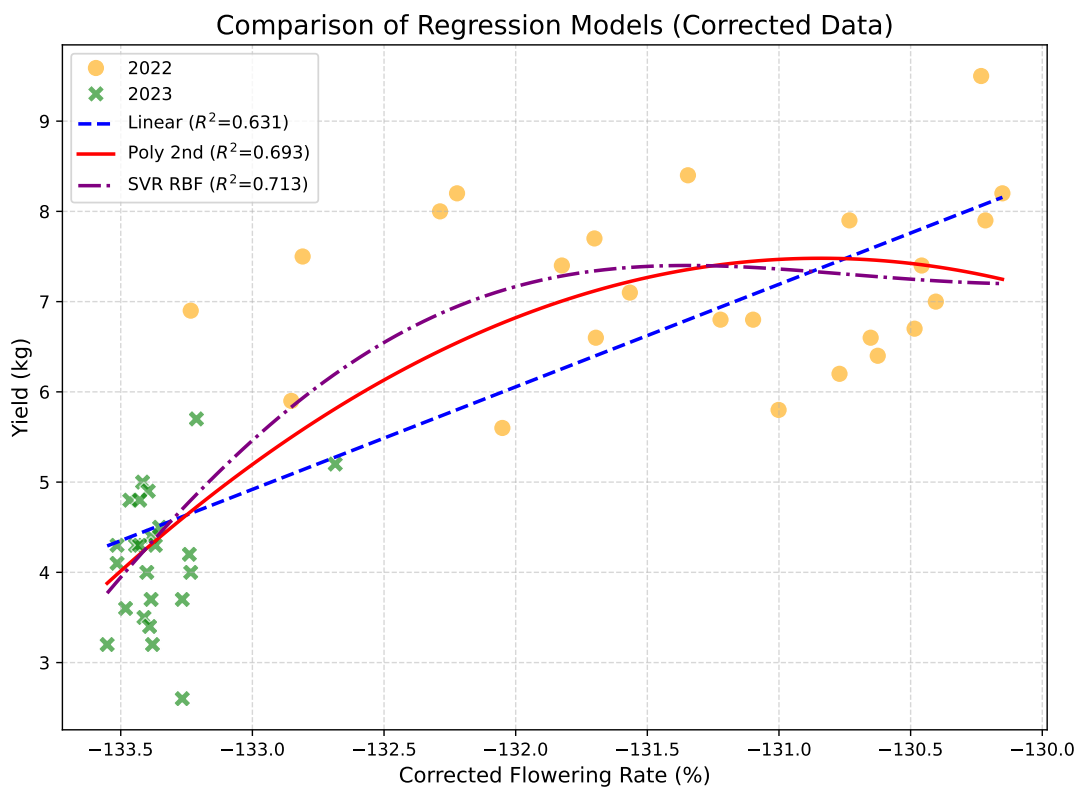
### 5.3.2 回帰分析による傾向の可視化

本実験の推定花量と実測収量の関係, および各回帰モデルによる手法ごとのフィッティング結果を図 5.4 に示す. いずれの手法においても, 推定花量と収量の間には負の相関関係が見られた. 線形回帰 (Linear) ではデータ全体の傾向を捉えているものの, 一部のデータ点での誤差が大きい. これに対し, SVR および二次多項式回帰では, データの分布に対してより柔軟に適合しており, これが  $R^2$  の向上に寄与したと考えられる.

### 5.3 推定花量に基づく収量予測モデルの評価実験結果



(a) U-Net



(b) RGB

図 5.4: 手法ごとの推定花量と実収量の回帰分析結果

## 5.4 考察

### 5.4.1 画像解析手法の比較と精度の要因

実験結果より、RGB 閾値法も比較的高い相関 ( $R^2 \approx 0.623 - 0.707$ ) を示した。これは、対象とした果樹園の撮影条件において、白い花と緑色の葉のコントラストが比較的明確であり、単純な色閾値処理でも花量の大きな傾向を捉えることが可能であったためと考えられる。

しかしながら、最高精度を達成したのは U-Net 法と SVR を組み合わせたモデル ( $R^2 = 0.714$ ) であった。RGB 法では、太陽光の反射や空などのノイズを花として誤検出するリスクが高いが、U-Net 法では形状や文脈を学習しているため、より純粋な花領域を抽出できている可能性がある。この特徴量の質の差が、特に SVR のような複雑な回帰モデルを用いた際に、わずかながらも精度の向上として現れたと推察される。

### 5.4.2 花量と収量の非線形性に関する考察

回帰モデルの比較において、線形回帰 (Linear) よりも二次多項式回帰や SVR の精度が高かったことは、果樹園芸学的に重要な示唆を与えている。単純な線形モデル ( $R^2 \approx 0.623 - 0.641$ ) に比べ、非線形モデルが高い適合度を示したことは、花量と収量の関係が単純な比例関係ではないことを意味する。

具体的には、着花量が少ない範囲では収量は花数に依存して増加するが、ある一定量を超えると、栄養競合による生理落果や果実肥大の抑制が生じ、収量の増加が頭打ちになる (飽和する)、あるいは複雑な挙動を示す可能性がある。SVR が最も高い精度 ( $0.707 - 0.714$ ) を記録したことは、このような生物学的な非線形性や、データに含まれる複雑な境界条件をモデルが適切に学習できた結果であると考えられる。

## 5.4 考察

### 5.4.3 推定花量が収量モデルに及ぼす影響

前節までの解析では、複数年（2022年，2023年）のデータを統合して回帰分析を行った。しかし、一般にカンキツ類果樹栽培においては、気象条件や隔年結果（Alternate Bearing）の影響により、年ごとの着花量および収量の平均的な水準が大きく変動する。そこで、本提案手法による推定花量と収量の相関が、樹体個々の生理的な着花能力、個体差に起因するものか、あるいは年ごとの隔年結果性に強く依存するものかを検証するため、年次別の標準化を用いた追加実験を行った。

具体的には、各年度  $y$  における樹体  $i$  のデータ（推定花量または実測収量） $x_{iy}$  に対し、その年度の平均値  $\mu_y$  および標準偏差  $\sigma_y$  を用いて、以下の式 (5.1) により標準化を行った。

$$z_{iy} = \frac{x_{iy} - \mu_y}{\sigma_y} \quad (5.1)$$

この処理により、豊作年・不作年といった年ごとのスケールの違いや平均的なバイアスを除去し、各年度内での相対的な位置関係のみを抽出して相関を再評価した。

#### 標準化データを用いた解析結果

推定花量および収量を年ごとに標準化したデータを用いた回帰分析の結果を図 5.5 に示す。補正後の生データを用いた全体解析（図??）では、 $R^2 > 0.6$  の比較的高い相関が確認されたが、年ごとの標準化を行ったデータにおいては、決定係数が著しく低下し、ほぼ無相関（ $R^2 \approx 0.02 - 0.06$ ）に近い結果となった。

これは、全体解析で見られた高い相関が、個々の樹体の着花量と収量の比例関係を純粹に示していたわけではなく、花が多い年は収量も多い（2022年）、花が少ない年は収量も少ない（2023年）という、年単位のクラスタ間の差異を強く反映していたことを示している。

## 5.4 考察

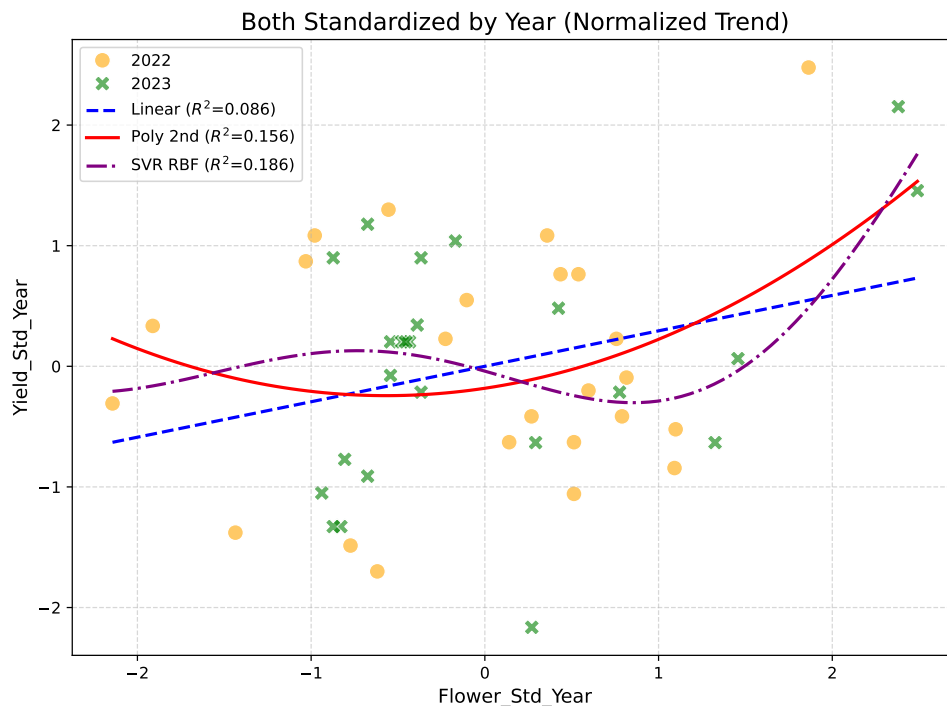


図 5.5: 推定花量及び収量を年ごとに標準化した際の回帰分析結果

### 収量決定要因における環境と管理の支配性

標準化データにおける相関の消失は、本対象圃場における文旦の収量変動において、個体レベルの初期着花数の差異よりも、他の支配的な要因が働いていることを示唆している。以下にその主要な2つの要因を考察する。

■(1) 年次変動 (Year Effect) の支配的影響 全体解析と標準化解析の乖離は、気象条件や隔年結果サイクルに起因する年次変動の影響が極めて大きいことを示している。すなわち、不作年においては、個体レベルで相対的に多量の花を着生させたとしても、豊作年の平均的な収量水準には到達し得ないという「環境的な限界」が存在することが定量的に確認された。

■(2) 栽培管理 (摘果) による平滑化 同一年度内において、花量と収量の相関が見られなかった (花が多くても実が多くなるとは限らない) 要因として、人為的な「摘果」の影響が強く推察される。実際の栽培現場では、樹勢維持と果実品質の均一化を目的として、着花数

## 5.4 考察

に関わらず最終的な着果数は「葉果比（葉数に対する果実数）」等の基準に基づき、一定範囲内に調整される。この管理作業により、開花時点での個体差（花の多さ）が収穫時には平準化されるため、画像解析が正確に花量を推定していたとしても、個体レベルでの最終収量との直接的な相関は低くなったと考えられる。

### 5.4.4 本システムの有用性の再定義

以上の結果より、本システムは個々の樹体の最終収量を厳密に予測するものではないことが明らかとなった。しかしながら、これは本システムの有用性を否定するものではない。

本システムは、その年の作柄（表年か裏年か）およびエリア全体の期待収量を、開花期という早期に判定するツールとして高い有用性を持つ。個体差の相関が低いことは、むしろ適切な摘果管理が行われている証左とも解釈できる。したがって、本システムを用いることで、生産者はその年に必要な摘果作業の労働量予測や、出荷資材・物流計画の早期立案といった、マクロな視点での栽培管理意思決定に貢献できると結論付けられる。

# 第6章

## 結論

本研究では、UAV 空撮画像と深層学習技術を統合し、露地栽培における土佐文旦の開花量を定量化することで、早期に収量を予測する手法を構築・評価した。

技術的な側面においては、YOLOv11 による樹体検出、SAM による高精度な樹冠抽出、および U-Net による花領域セグメンテーションを階層的に組み合わせる提案手法が有効であることを示した。比較実験の結果、提案手法は従来の色情報に基づく閾値処理手法と比較して、背景ノイズや照明変動に対して堅牢であり、高い IoU および AIC 基準でのモデル適合度を達成した。これにより、複雑な露地環境下においても、画像解析によって「花量」を客観的かつ高精度に数値化可能であることが実証された。

農業的な側面における収量予測の検証では、推定された花量と実収量との間に、 $R^2 > 0.7$  という高い相関関係が確認された。さらに、年次ごとの標準化を用いた詳細な解析により、この相関は樹体個々の生理的特性よりも、気象条件や隔年結果サイクルに起因する「年次変動（表年・裏年）」に強く依存していることが明らかとなった。個体レベルでの相関が希薄であったことは、現場において適切な「摘果」管理が行われ、着花数の多寡に関わらず最終的な着果数が適正範囲に平準化されている実態を裏付けるものである。

以上の結果より、本システムは個々の樹体の最終収量をピンポイントで予測するものではないものの、開花期という極めて早期の段階において、その年の作柄傾向（豊作・不作の水準）やエリア全体の期待収量を判定するツールとして極めて有用であると結論付けられる。本技術の実装により、生産者は摘果に必要な労働力の適正配置や、出荷資材・物流計画の早期立案が可能となり、データ駆動型の効率的な果樹経営への寄与が期待される。

今後の展望として、より長期間にわたるデータの蓄積による年次変動モデルの精緻化や、

摘果後の着果数計測との組み合わせによる予測精度の向上が挙げられる。また、提案手法に入力するデータセットを着果樹体に置き換えるとさらに高精度での収量予測が可能になると考える。撮影条件についても天候、撮影高度を変更して予測に適した条件を見つける必要がある。

# 謝辞

本研究を行うにあたって、高知工科大学情報学群吉田真一教授にはご指導いただき心から感謝いたします。また、本研究の遂行にあたり、貴重な実験データの提供ならびに多大なるご協力をいただきました、高知県農業センターの皆様にも深く感謝いたします。特に、谷本祐氏には、データの詳細に関するご教示や、本研究に対する有益なご助言をいただきました。ここに記して深く謝意を表します。さらに、本研究の副査をお受け頂きました高知工科大学情報学群岩田誠教授並びに、福本昌弘教授にも深く感謝いたします。

# 参考文献

- [1] 農林水産省, “農業労働力に関する統計,” <https://www.maff.go.jp/j/tokei/sihyo/data/08.html>, 2026年1月28日.
- [2] 農林水産省, “スマート農業をめぐる情勢について,” <https://www.maff.go.jp/j/kanbo/smart/attach/pdf/index-297.pdf>, 2026年1月28日.
- [3] 高知大学 IoP 共創センター, “IoP (Internet of Plants) とは,” <https://www.kochi-u.ac.jp/iopc3/about.html>, 2026年1月28日. p.3, 2021.
- [4] カンキツの調査方法編集委員会 (編), “カンキツの調査方法,” 果樹試験場興津支場研究資料, No. 6, 1987.
- [5] 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.
- [6] M. Kitano, K. Nomura, T. Yamazaki, T. Iwao, M. Saitou, M. Mori, D. Yasutake, T. Kaneko, H. Ukeda, S. Ishizuka, T. Fujiwara, and T. Okabayashi, “Internet of Plants (IoP) empowers bottom-up innovations in greenhouse horticulture,” *Environmental Control in Biology*, vol. 60, no. 1, pp. 3–12, 2022.
- [7] 北野雅治, “Hybrid AI モデル群による作物群落の生理生態情報の営農現場での可視化,” 科学研究費助成事業 (基盤研究 (B)) 研究成果報告書, 課題番号 20H03131, 2023.
- [8] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [9] A. Koirala, K. B. Walsh, Z. Wang, and C. McCarthy, “Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of ‘MangoYOLO’,” *Precision Agriculture*, vol. 20, pp. 1107–1135, 2019.
- [10] P. A. Dias, A. Tabb, and H. Medeiros, “Apple flower detection using deep convo-

## 参考文献

- lutional networks,” *Computers in Industry*, vol. 99, pp. 17–28, 2018.
- [11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *Proc. CVPR*, pp. 779–788, 2016.
- [12] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLO,” version 11.0, <https://github.com/ultralytics/ultralytics>, 2024.
- [13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proc. CVPR*, pp. 580–587, 2014.
- [14] 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.
- [15] O. Ronneberger et al., “U-net: Convolutional networks for biomedical image segmentation,” *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 234–241, 2015.
- [16] A. Kirillov, E. Mintun, N. Ravi, et al., “Segment Anything,” *arXiv preprint arXiv:2304.02643*, 2023.
- [17] N. Ravi, et al., “SAM 2: Segment Anything in Images and Videos,” *arXiv preprint arXiv:2408.00714*, 2024.
- [18] C. Ryali, Y.-T. Hu, D. Bolya, et al., “Hiera: A Hierarchical Vision Transformer without the Bells-and-Whistles,” *International Conference on Machine Learning (ICML)*, pp. 29430–29446, 2023.
- [19] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [20] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. CVPR*, vol. 1, pp. 886–893, 2005.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. CVPR*, pp. 770–778, 2016.

## 参考文献

- [22] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems (NIPS)*, vol. 25, pp. 1097–1105, 2012.