

## 原著論文

ドローン観測によるキャベツ生育ステージ  
推定学習モデルの提案

朴 鍾杰\*・張 九勻\*

**要旨：**本研究では、ドローン観測画像を用いてキャベツの成長ステージを推定するために最適な学習データ生成方法について調べた。主なアプローチとして、1. 観測高度に基づくキャベツ成長ステージの推定、2. 混合解像度モデルを用いた成長ステージの推定について行った。物体検出には、物体の位置とカテゴリを同時に特定するYOLOモデルを採用した。学習データは、高度10mから得られた画像を使用して学習モデルを構築した場合、高度10mの検証データに対するmAP@0.5は0.94であったが、100mの高度での検証データに対する精度は0.156と非常に低い結果となった。逆に、高度100mで収集した画像から学習したモデルは、高解像度検証画像で最も低い精度を示した。そのため、混合解像度学習データを使用して学習した結果、高解像度画像と低解像度画像の両方において高い検出精度が得られることが分かった。

**キーワード：**キャベツ, 生育ステージ, 混合解像度モデル, YOLO, ドローン

Learning Model to Estimate Cabbage Growth Stages through Drone  
Observation

Jonggeol PARK\* and JiuYun ZHANG\*

**Abstract:** In this study, we evaluated the accuracy improvement in machine learning used for estimating the growth stages of cabbage using drone images captured at different altitudes. We adopted the YOLO model for object detection, which simultaneously identifies the position and category of objects. We constructed the learning model using images obtained at an altitude of 10 meters and the mAP@0.5 for the validation data at this altitude was found to be 0.94. However, the accuracy for validation data at an altitude of 100 meters was significantly lower at 0.156. Moreover, the model trained with images collected at an altitude of 100 meters exhibited the lowest accuracy, especially for high-resolution images. We found that high detection accuracy was achieved for both high-resolution and low-resolution images by using mixed-resolution training data.

**Keywords:** Cabbage, Growth Stage, Mixed-resolution Model, YOLO, Drone

## 1. はじめに

日本の主要な農業労働者（15歳以上で世帯員のうち、ふだん仕事として主に自営農業に従事している者）は、平成17年（2005年）224万人年から、令和2年（2020年）には136万人に40%減少した。また、農業従事者の平均年齢も64.2歳から67.8歳と年々上昇している [1]。農業分野では、高齢化の進行による労働力不足が深刻な課題となっている。これに対処するため、スマート農業が提案され、ロボット技術や情報通信技術を活用して、労力削減、精密農業、高品質生産などを試みている。ただし、スマート農業技術を従来の方法に適用しても生産性向上には限界があるため、自動収穫ロボットなどのスマート農業機械を最大限に活用できる農業現場の変革が重要とされてる [2]。また、農林水産省は、人口減少にもかかわらず生産水準を維持できる高い生産性を持つ食料供給システムを確立するため、スマートフードチェーンプラットフォームを活用して、生産、流通、販売の一体化を検討している。

キャベツは、農地の広い領域で栽培され、農場の規模が大規模化している野菜の一つである。特に千葉県は全国的にキャベツの生産が盛んで、その生産量は愛知県や群馬県に次いで多い地域の一つである。ところが、農林水産省の統計データによれば、過去10年間（2012年から2021年まで）において、日本全体でのキャベツの生産量は2.9%増加したが、千葉県ではキャベツの作付面積が9%減少し、その結果、収穫量も7%減少した。キャベツは労働集約型の作物であり、収穫作業の省力化が不可欠である。同じ日に種まきを行っても、生育環境によって株ごとの成長が異なり、出荷日が異なるため、収穫時期において一斉収穫が難しく、通常は人手による収穫が行われる。キャベツなどの作物は機械収穫が可能だが、現行の方法では収穫基準に達していない株も同時に収穫され、出荷量が減少してしまうため、手作業に頼っている。キャベツ農家の収益性を向上させるためには、スマート農業による成長モニタリングや収穫自動化など、自動化が不可欠である。

ドローン技術は急速に進化し、その発展に伴い農業分野でも広く活用された。ドローンリモートセンシングを用いることで、低コストで小規模な領域で高解像度の画像を取得でき、また、マルチスペクト

ルカメラを搭載することで、作物の水ストレス、クロロフィル指数、光合成有効放射吸収能 (fAPAR)、窒素含有量などに関する情報を迅速に取得できる [2]。ドローンは飛行高度を調整できるため、高い飛行高度では広範囲を短時間で撮影できるが、地上の解像度は低いため小さな作物を識別するのが困難である。一方、飛行高度を低くすると、作物の葉の形状まで識別できるが、観測に時間がかかる。田中ら (2021) は、キャベツを株単位で効果的に認識するための最適なドローンの飛行高度（地上の解像度）について調査し、その結果、高度30m（地上解像度1.3cm/px）では、広範囲を短時間で観測し、キャベツの株を抽出するのに有効であることを示した。

機械学習を用いた物体認識では、大規模なデータセットを使用してモデルを構築し、正解と予測の誤差を最小化するための処理が行われる。そのため、ユーザは入力データ内で出力精度に寄与する有用な特徴を直接特定することは難しい。機械学習モデルにおいて有効な特徴を調べる方法として、特徴量選択 (Feature Selection) と特徴量抽出 (Feature Extraction) の2つのアプローチがある。特徴量選択は、既存の特徴量セットから最も有用な特徴量のサブセットを選び出し、モデルの性能を向上させ、計算コストを削減する方法である。主にフィルター法 (Filter Method)、ラッパー法 (Wrapper Method)、および埋め込み法 (Embedded Method) がある。特徴量抽出は、元の特徴量セットから新しい特徴量を生成するプロセスであり、主に主成分分析 (Principal Component Analysis, PCA)、独立成分分析 (Independent Component Analysis, ICA) などがある。畳み込みニューラルネットワーク (Convolution Neural Network, CNN) を使用した物体検出においても、有効な検出領域を特定するための研究が行われた。Zhouら (2015) は、Grad-CAMを用いてCNNの内部特徴を可視化し、物体の位置を特定する方法を提案した。CNNモデルが物体検出やセグメンテーションにおいて、どの画像領域に注目して判断するかを可視化することで、モデルの透明性を高め、モデルの内部動作を理解しやすくするツールとして幅広く活用された (Jooら、2019年、Panwarら、2020年、Moujahidら、2022年)。

本研究は、ドローンの飛行高度によって異なった地上解像度の観測データを用いても、キャベツの成長ステージを精度よく推定できる最適な学習データ

の作成方法について調べる。まず、高解像度観測データを用いて10mから100mまで10mごとに変えながらサンプリング画像を作成し、それぞれの高度の学習データを用いてキャベツの成長ステージの検出精度検証を行う。また、本研究で提案する10mから100mまですべての高度の学習データを用いた混合解像度モデルによる高度別キャベツの成長ステージの検出を行う。最後に2つの方法の検出精度比較を行う。

## 2. データ

### 2.1 観測地

本研究の観測対象地は、千葉県千葉市下田町に位置する吉田農園のキャベツ畑である(図1)。ドローンによる観測は、キャベツの苗を植えた定植期(2021年8月30日および9月11日)、生育期(9月20日および9月27日)、および結球収穫期(10月6日および10月15日)の計6回行われた。ドローンを用

いた圃場の観測には、合計で27分を要し、航路を2,134m飛行しながら、約470枚の画像を撮影した。図1の赤い枠で示された領域にはキャベツが栽培されており、その他の場所にはブロッコリー、ショウガ、サツマイモなどの異なる作物が栽培されている。

### 2.2 ドローン仕様と飛行高度による解像度

本研究では、DJI P4 Multispectral (P4M) を使用した。P4Mは、マルチスペクトルセンサー(青、緑、赤、レッドエッジ、近赤外)および可視光カメラを搭載している(表1)。P4Mの重量はバッテリーを含めて1,487gであり、最大飛行時間は27分である。P4Mの飛行は、DJI GS PROアプリケーションを使用して、撮影範囲と撮影高度を設定することで、最適な飛行経路が自動的に生成され、その経路に従って撮影が行われる。

表2に自動飛行設定のパラメータを示す。ドローンの飛行は、国土交通省の無人飛行機(ドローンやラジコン機など)の飛行規則に従って行われた。飛

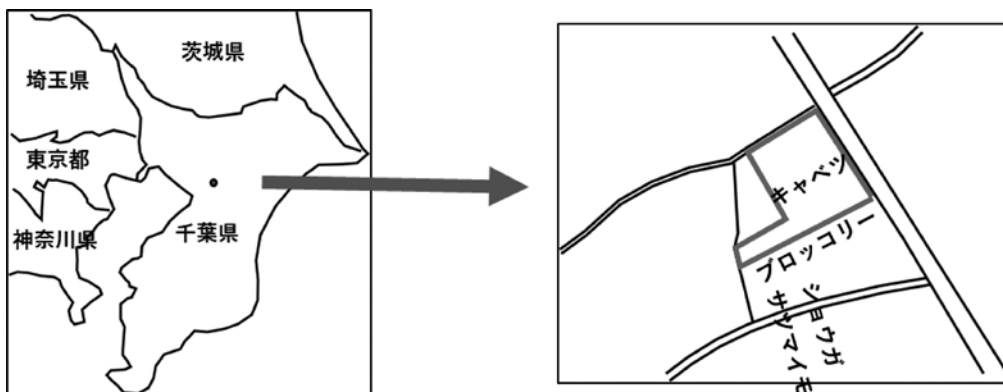


図1 観測地

表1 Phantom 4 Multispectralの仕様

センサー	6個の1/2.9インチCMOS(可視光RGBセンサー1台、マルチスペクトルイメージング用モノクロセンサー5台を含む) 各センサー:有効画素数2.08MP	
フィルター	ブルー(B)	450nm±16nm
	グリーン(G)	560nm±16nm
	レッド(R)	650nm±16nm
	レッドエッジ(RE)	730nm±16nm
	近赤外(NIR)	840nm±26nm
レンズ	FOV(視野角):62.7° 焦点距離:5.74mm(35mm版換算:40mm) 絞り:F2.2	
地上解像度	(H/18.9)cm/pixel. Hは機体高度(単位:m)	

表2 自律飛行設定パラメータ

飛行高度	10m
地上分解能 (GSD)	0.53cm/pixel
飛行時間	約20分
撮影モード	同時間隔 (2秒ごと)
航路上のオーバーラップ率	65%
航路間のオーバーラップ率	65%

行高度に応じた地上解像度は、高度10mの場合の画像の地上解像度が0.53cm/pixelであり、高度100mの場合は5.29cm/pixelである。

### 3. 方法

#### 3.1 YOLOv5 (You Only Look Once Version 5) のデータ拡張とアーキテクチャ

YOLOはリアルタイム物体検出を行うための効率的なディープラーニングアルゴリズムであり、物体の位置とカテゴリを同時に予測できる。YOLOは、大規模なラベル付きデータセットを使用してトレーニングされ、高い精度で物体検出を実現する学

習済みモデルである。YOLOv5のアーキテクチャの特徴の一つはFeature Pyramid Network (FPN) である。FPNは異なる解像度の特徴マップを結合し、物体の位置やスケールに対する感受性を向上させた。これにより、小さな物体から大きな物体まで効率的に検出できるようになった。YOLOv5x6モデルでは、最終予測層 (23層、26層、29層と32層) で4つの異なるスケールの特徴マップを使用して予測を行う (図2)。

機械学習、特にディープラーニングによる画像認識では、大量の学習データが必要である。しかし、しばしば十分なデータを収集することが難しかった

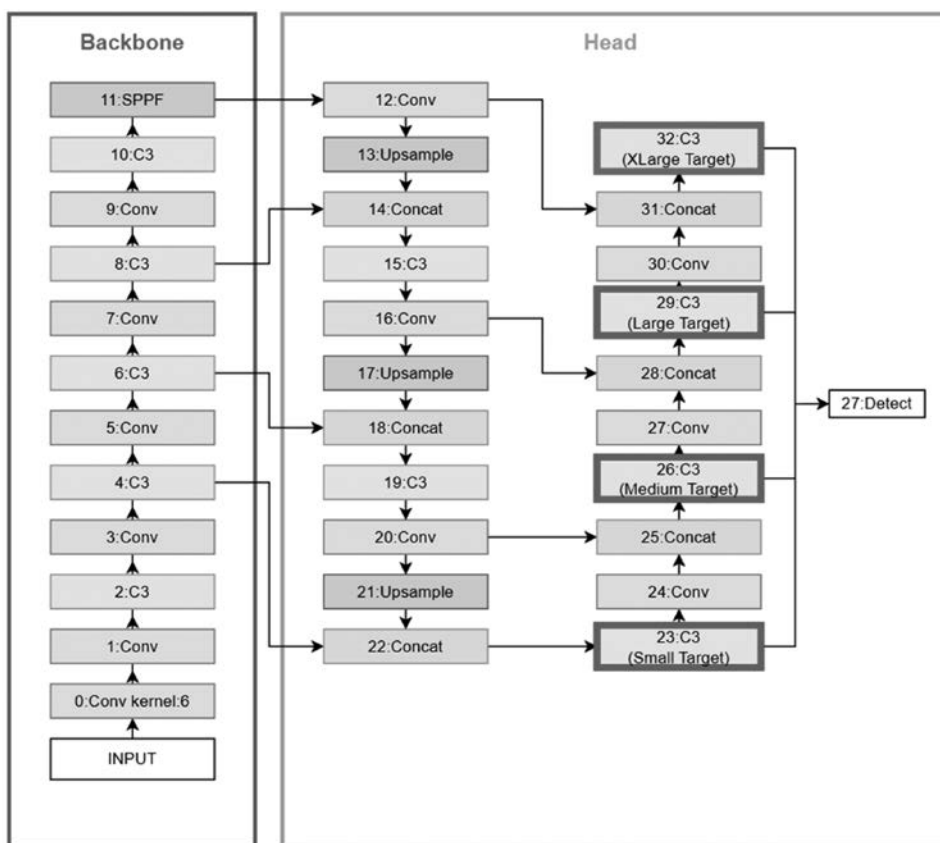


図2 YOLOv5x6モデルのアーキテクチャ

め、データ拡張 (Data Augmentation) の手法を使用して、学習データの多様性を増やすことが一般的である。本研究では、YOLOv5の基本データ拡張に加えて、パラメータを画像回転に45 (+/- deg)、画像の拡大縮小に0.5 (+/- gain) を設定し、学習を行った。

### 3.2 学習モデルの検証

学習モデルの精度評価指標には、精度 (Accuracy)、適合度 (Precision)、および再現度 (Recall) がある。学習モデルによって検証した結果 (予測値と真値) を用いて混同行列を作成する。混同行列の対角成分は正しく分類された結果を示す。Accuracyはすべてのデータのうち、正しく分類されたデータ (対角成分) の割合で、モデルの全体的な性能を示す指標である (式1)。しかし、データが不均衡な場合、Accuracyだけではモデルの評価が難しいことがある。そのため、Precision、Recall、およびF値を参照して総合的な評価を行う。Precisionは、モデルが正と予測したデータのうち、実際に正しかった割合を示す指標である (式2)。ただし、Precisionは偽陰性 (FN) を考慮しない。これを補完するためにRecallが用いる。Recallは実際に正であるデータのうち、モデルが正しく予測した割合を示す指標である (式3)。PrecisionとRecallはトレードオフの関係

にあり、一方の値が高い場合、もう一方の値は低くなる。F値はPrecisionとRecallの両方を考慮してモデルの性能を評価する指標である (式4)。F値はPrecisionとRecallのどちらを重要視するかに応じて異なる計算方法があり、高いF値はPrecisionとRecallのバランスが取れた性能を示す。

YOLOモデルは信頼度スコア (Confidence score) を使用して、物体の検出と同時に識別を行うアルゴリズムである。Confidence scoreの閾値を低く設定すると、予測枠 (画像領域内で物体が存在する可能性のある領域) の数が増え、Recallは向上するが、同時に誤検出が増え、Precisionは低下する。逆に、閾値を高く設定すると、高信頼度の予測枠のみが選択され、誤検出は減少するが、同時に本物の物体の検出漏れが増え、Precisionは向上し、Recallは低下する。したがって、モデルの評価時には、信頼度スコアの閾値に対するPrecisionとRecallの変化を総合的に比較する必要がある。Confidence scoreの閾値の変化に対するPrecisionとRecallの値をプロットし、その曲線下の面積 (Average Precision : AP) を計算してモデルの性能を評価する。AP値が1に近いほど、モデルの性能が優れていることを示す。物体検出モデルの最も一般的な評価指数はmAP@0.5 (Mean Average Precision at IoU 0.5) である。mAP@0.5は、

表3 2値分類における混同行列と評価指数

	正 (真値)	負 (真値)
正 (予測)	TP	FP
負 (予測)	FN	TN

TP : True Positive。実測値と予測値の両方が正であったもの。

FP : False Positive。実測値が負なのに、誤って正と予測値したもの (誤検知、偽陽性)。

TN : True Negative。実測値と予測値の両方が負であったもの。

FN : False Negative。実測値が正なのに、誤って負と予測値したもの (偽陰性)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{式1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{式2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad \text{式3}$$

$$F1 = \frac{2\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{式4}$$

予測枠と正解枠のIoU (Intersection of Union) が50%以上の場合に正しいと判断した場合の平均AP値である。

## 4. 結果

### 4.1 学習データセット作成と学習

1回のドローン観測で得られた470枚のRGB画像

(1,600×1,300)をDJI TERRAソフトウェアを使用してオルソモザイク画像(10,389×13,256)を生成した。機械学習の学習データと検証データセットを作成するために、オルソモザイク画像を300×300ピクセルサイズ(サブセット画像)に分割した。各観測日のサブセット画像から、キャベツを含む80枚のポジティブサンプルと、キャベツを含まない20枚の

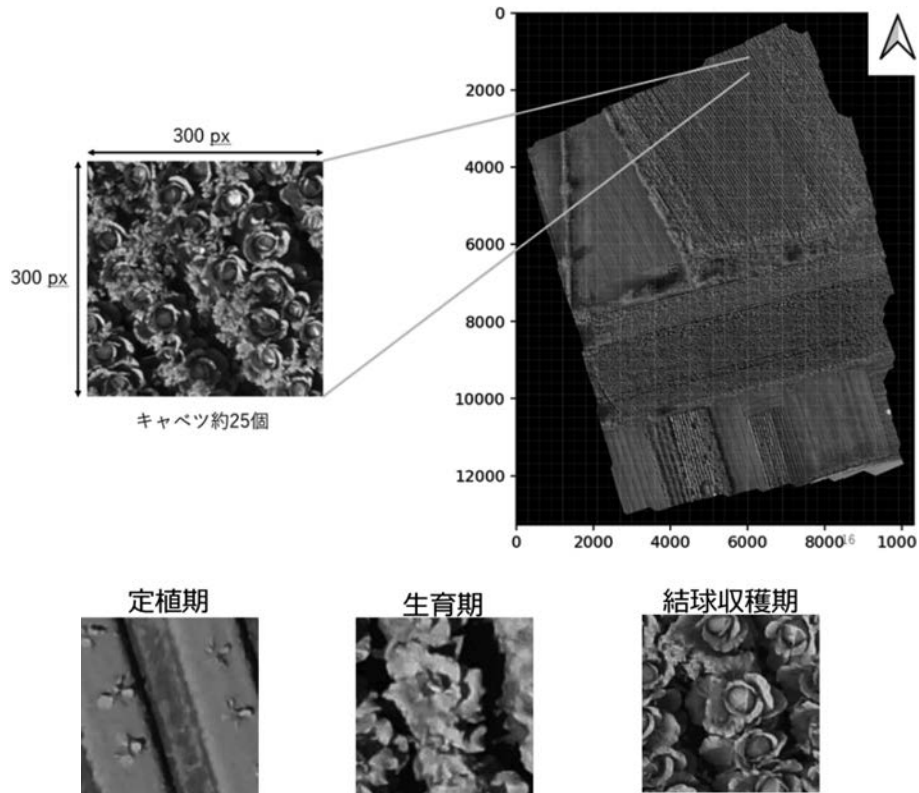


図3 オルソモザイク画像から300×300サイズのサブセット画像

表4 生育ステージに合わせた学習データ(学習用と検証用)数

生育ステージ	Train	Validation	Total
定植期	3,241	809	4,050
生育期	2,328	567	2,895
結球収穫期	2,839	716	3,555
Total	8,408	2,092	10,500

表5 YOLOv5学習パラメータ

使用するデータ	300×300画像
Epochs	100
Batch-size	2
使用する学習モデル	Yolov5x6.pt
Image-size	640

ネガティブサンプルを選んでラベリングを行った。ポジティブサンプル画像には約25個のキャベツが含まれている。ラベリングはキャベツの成長段階に合わせて、定植期（定植活着期）、生育期（葉枚数展開期）、結球収穫期の3つの段階に分けて行った。図3は、9月27日のオルソモザイク画像から切り取った300×300ピクセルの例である。ラベリングは定植期に4,050個、生育期に2,895個、結球収穫期に3,555個を選び、その中で80%を学習データ、残り20%を検証データとして用いた（表4）。表5は、YOLOv5の学習時のパラメータである。

#### 4.2 学習データによるキャベツ生育ステージ識別のための最適モデル決定

物体検出における学習モデル検証は、予測枠と正解枠の交差率が50%以上（IoU $\geq$ 50%）の場合正解（mAP@0.5）とする評価指数を多く用いる。しかし、IoU値は任意で決定できるため、交差率を95%（mAP@0.95）にすると検出精度が低くなる。そこで、各生育ステージに対してIoU値を0.5から0.05刻みで0.95まで（mAP@[0.5:0.05:0.95]）変えながら最大mAP時のIoU値を調べた。その結果、定植期のIoU値は0.547、生育期では0.425、結球収穫期では0.543、すべてのステージを含めた場合0.505であった（表6）。本研究ではすべてのステージに安定した結果を求めるためにIoU値を0.5とした学習

モデルの重みデータをモデルに読み込んだ。学習時のmAP@0.5の精度（Precision、Recall）を表6に示す。図4はmAP@0.5によって得られた生育ステージ別検出結果画像である。定植期、生育期と結球収穫期のすべてのキャベツに対して正しく検出できたことが分かる。

#### 4.3 高度別キャベツ生育ステージ推定

ドローン観測は高度10m（0.529cm/pixel）で実施したが、画像リサンプリング手法を使用して、観測高度を20m（1.06cm/pixel）から100m（5.29cm/pixel）までの10m間隔での画像を生成し、高度別学習モデルを作った。学習モデル作成時の設定は入力画像の多様性を増やすために画像拡張法を適用した。高度別ラベリングは高度10mのラベリングデータをそのまま適用した。

まず、高度10mで訓練したモデルを使用して、高度ごとのmAP@0.5を求めた（図5）。その結果、高度10mでは0.94であったが、高度40mでは0.63、高度100mでは0.156など、高度が上がるにつれて精度が低下することがわかった。図6は、高度ごとの学習モデルを用いて、高度10mから100mまでの検証データの成長ステージ判定を行った結果である。共通的な特徴は学習モデルの高度より検証データの高度が高くなればなるほど精度が低くなることである。高度10m学習モデルの場合、高度40m以上の検

表6 各ステージにおけるmAP@0.5モデルを用いた場合の評価指数

生育ステージ	Precision	Recall	mAP@0.5	mAP@[0.5:0.95]
定植期	0.932	0.959	0.969	0.547
生育期	0.849	0.871	0.877	0.425
結球収穫期	0.849	0.871	0.877	0.543
すべてのデータ	0.905	0.935	0.94	0.505

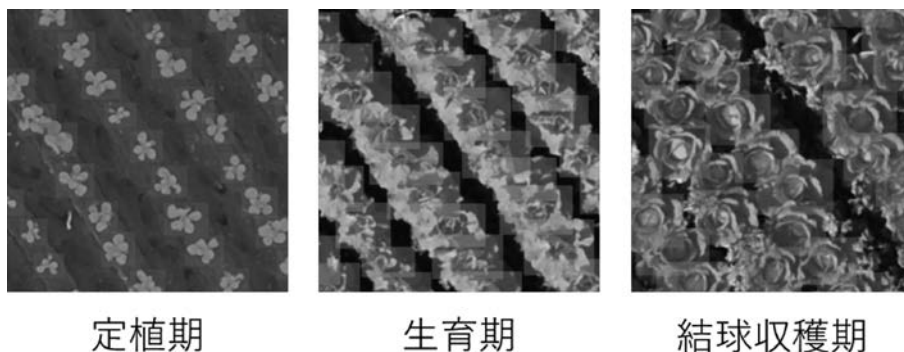


図4 学習データにおけるmAP@0.5モデルによる生育ステージ検出結果

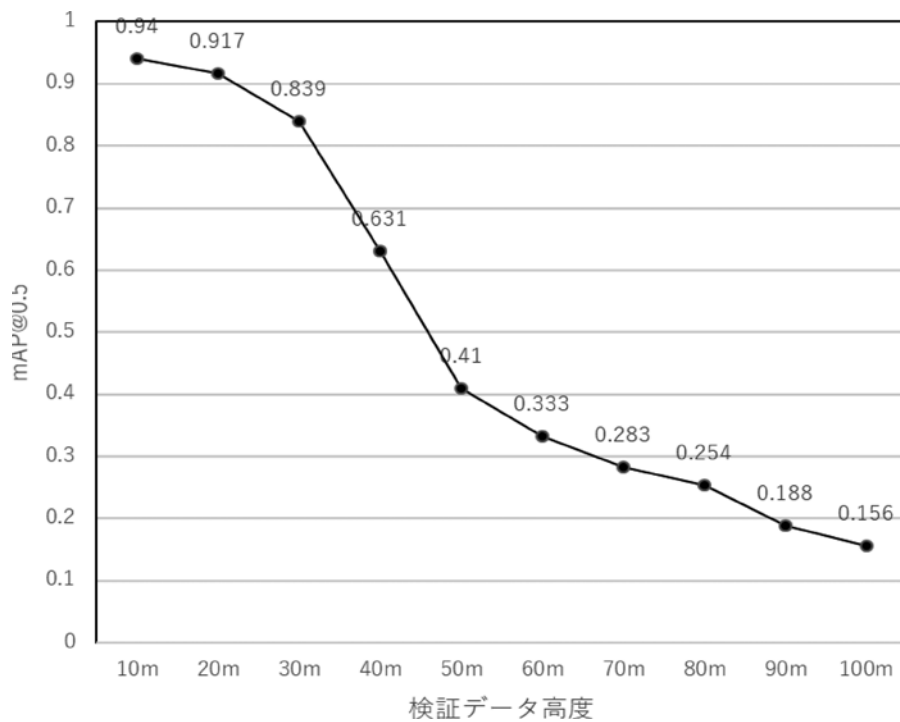


図5 高度10m学習データモデルを用いた場合の高度別mAP@0.5精度

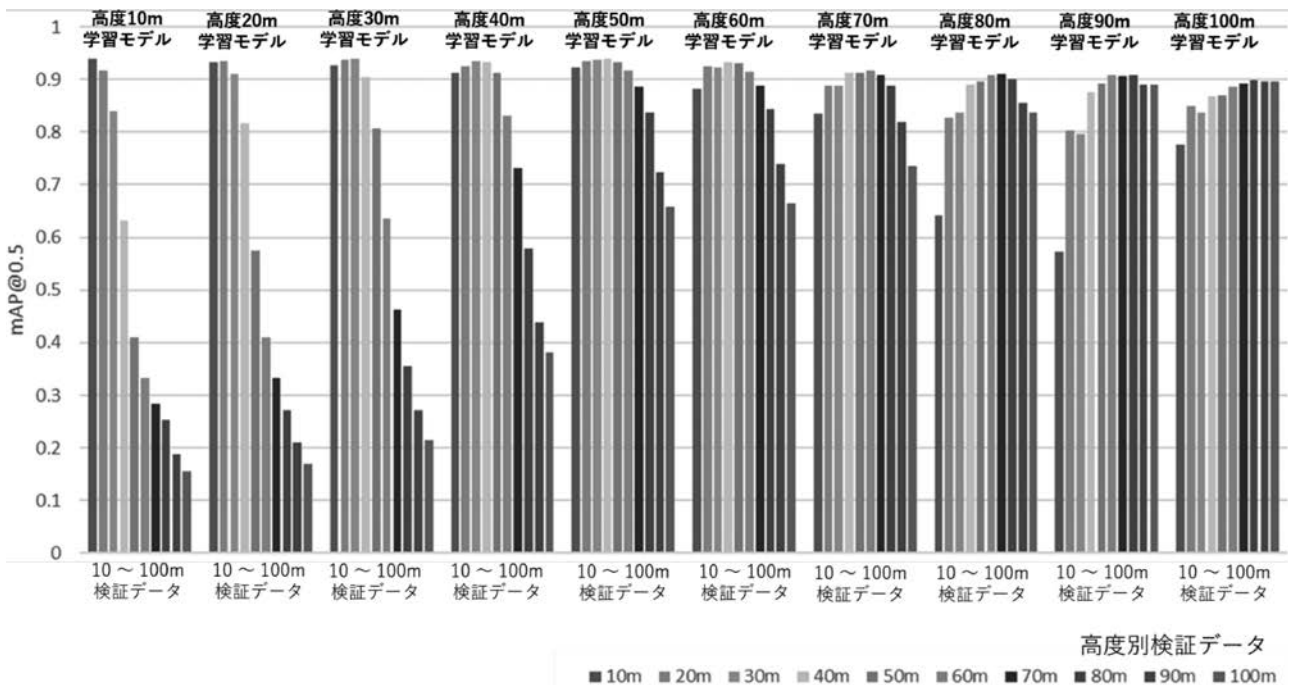


図6 高度別学習モデルによる高度別キャベツ生育ステージ検出精度 (mAP@0.5)

証データに対してmAP@0.5は0.7以下である。高度30m学習モデルの場合、高度60m以上の検証データに対してmAP@0.5が0.7以下であった。学習モデル高度より30m高くなると検証データに対する精度は急に低くなる。しかし、学習モデルの高度より低い高度の検証データに対しする精度の低下は大きくな

い。高度50mの学習モデルによる高度10mの検証データのmAP@0.5は0.88であり、高度80mの学習モデルによる高度20mの検証データのmAP@0.5は0.83であった。

#### 4.4 混合解像度モデルの提案

YOLOv5のデータ拡張のみでは、学習データの高



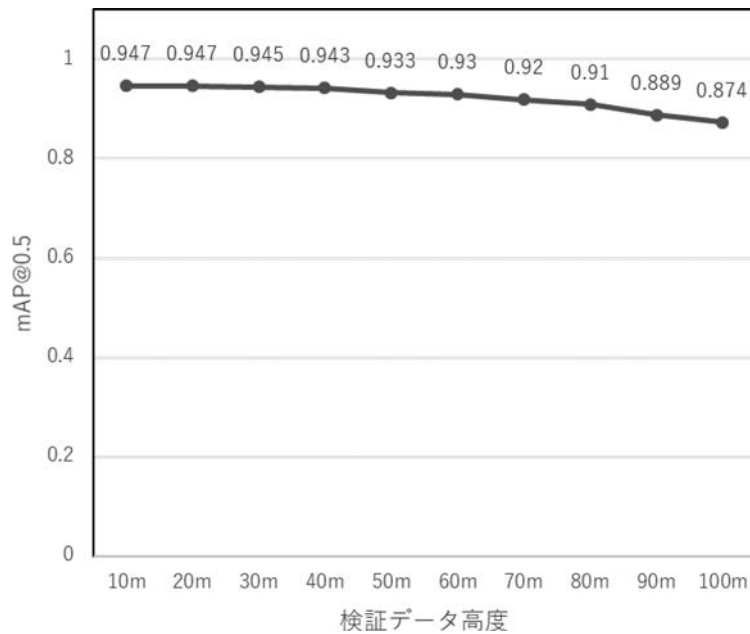


図7 混合解像度学習モデルを用いた高度別mAP@0.5

度と異なるデータを使用した場合、精度が低下することが明らかになった。したがって、本研究では、すべての高度にわたる学習データ（4.3節の10mから100mまでの高度別ラベリングデータすべて）を使用して学習モデル（混合解像度モデル）を構築した。学習時の設定として画像拡張法を用いた。図7は混合解像度モデルを使用した場合の高度ごとのmAP@0.5である。x軸は高度べつ（10mから100mまで）検証データである。高度10mの検証データに対する精度は0.947であり、高度100mの検証データに対する精度は0.874であった。混合解像度モデルを使用することで、異なる高度の検証データに対しても高い精度でキャベツの成長ステージを識別することが分かった。

## 5. まとめ

本研究では、異なる高度で撮影したドローン画像を用いて、キャベツの成長ステージを推定するための機械学習の精度向上について調べた。物体検出には、物体の位置とカテゴリを同時に特定するYOLOモデルを採用した。学習データは、高度10mから得られた画像を使用して学習モデルを構築した場合、高度10mの検証データに対するmAP@0.5は0.94であったが、100mの高度での検証データに対する精度は0.156と非常に低い結果となった。逆に、高度100mの画像から学習したモデルは、低高度で最も

低い精度を示した。そこで、ドローンの飛行高度によってキャベツ成長ステージ推定精度に影響されにくい混合解像度モデルを提案した。その結果、低高度画像から高高度画像データまで高い検出精度を得ることができた。

今後の課題としては、高度ごとの学習モデルと混合解像度学習モデルが検証データに着目する場所の特徴と違いに関して調べ、なぜ精度の違いが生じるのかを明らかにする必要がある。

## 謝 辞

本研究はR3年度からR5年まで「東京農業大学と東京情報大学との共同研究プロジェクト」として選ばれ、“Society5.0社会におけるレジリエンス農業確立に向けた多様な人々の能力を發揮するロボティク・プロセス・オートメーション（RPA）実装の加速化”の研究テーマの一環として行いました。ここで深い感謝の意を表明します。また、有益な助言・コメントを頂いた研究チームの先生方に感謝申し上げます。

## 引用文献

- [1] 青木 循. (2013). 新型キャベツ収穫機. 農業機械学会誌, 239-241.
- [2] 井上吉雄, 横山正樹. (2017). ドローンリモートセンシングによる作物・農地診断情報計測とそのス

マート農業への応用. 日本リモートセンシング学会誌37巻3号, 224-235.

- [ 3 ] 田中美咲, 濱侃, 鶴崎 幸, 柴戸靖志. (2021). キャベツほ場の生育を株単位で把握するためのドローン空撮方法と画像解析方法. 日本リモートセンシング学会誌, 41 (3), pp.375-385.
- [ 4 ] Joo, H. and Kim, K., 2019, Visualization of Deep Reinforcement Learning using Grad-CAM: How AI Plays Atari Games, IEEE Conference on Games (CoG), DOI: 10.1109/CIG.2019.8847950
- [ 5 ] Moujahid, H., B., Cherradi, Sarem, M. A., Bahatti, L., Eljialy, A., Alsaedi, A. and Saeed, F., 2022, Combining CNN and Grad-Cam for COVID-19 Disease Prediction and Visual Explanation, Intelligent Automation & Soft Computing, DOI: 10.32604/iasc.2022.022179
- [ 6 ] Panwar, H., Gupta, P. K., Siddiqui, M. K., Menendez, R. M., Bhardwaj, P. and Singh, V., 2020, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images, Chaos, Solitons & Fractals, Volume 140, November 2020, <https://doi.org/10.1016/j.chaos.2020.110190>
- [ 7 ] Zhang, J., Wang, X., Liu, J., Zhang, D., Lu, Y., Zhou, Y., Sun, L., Hou, S., Fan, X., Shen, S. and Zhao, J., 2022, Multispectral Drone Imagery and SRGAN for Rapid Phenotypic Mapping of Individual Chinese Cabbage Plants, DOI: 10.34133/plantphenomics.0007
- [ 8 ] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. and Torralba, A., 2015, Learning Deep Features for Discriminative Localization, arXiv:1512.04150v1 [cs.CV] 14 Dec 2015

#### Webサイト

農林水産省：農業農動力に関する統計、<https://www.maff.go.jp/j/tokei/sihyo/data/08.html> (2023年10月15日閲覧)

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-121.pdf> (2023年10月15日閲覧)

YOLOv5 Github：Ultralytics YOLOv5, <https://github.com/ultralytics/yolov5>, (2023年10月15日閲覧)