

## 短 報 (Short Communication)

# トドマツの樹冠画像からの球果検出に基づく 物体検出アルゴリズム YOLO のモデル比較

花岡 創<sup>1)\*</sup>、武津 英太郎<sup>2)</sup>

### 要旨

トドマツの樹冠画像から球果を検出する精度を、物体検出アルゴリズムの You Only Look Once (YOLO) の複数のモデルで比較した。YOLOv4 と YOLOv5m の検証では、両者とも 0.9 程度の average precision (AP) を示した。また、YOLOv5 にはニューラルネットワーク (NN) のレイヤーサイズが異なる 5 種類のモデル (YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x) があるため、これらの精度比較も行った。YOLOv5n を除く全てのモデルで AP が 0.9 となり、球果検出に関しては、NN のレイヤーサイズによらず同様の精度となった。

キーワード：YOLOv4、YOLOv5、物体検出、深層学習、着果評価

### 1. はじめに

近年、植物の状態管理や作物の自動収穫装置の実現等を目的として、物体検出 (object detection) 技術を応用した多様な技術開発が行われている。例えば、イチゴの成熟度合いの評価 (Fan et al. 2022) や、トマトの病害箇所の抽出 (Liu and Wang 2020)、果樹についての果実の検出 (Parico and Ahamed 2021, Itakura et al. 2021, Chen et al. 2021) などの事例がある。物体検出アルゴリズムにも様々なものが提案されているが、広く利用されているアルゴリズムの一つに You Only Look Once (YOLO) がある。上述の研究も全て YOLO を活用した事例である。花岡 (2021) は、北海道の主要造林樹種であるトドマツの着果量評価を効率化することを目的に、UAV を用いて撮影したトドマツの樹冠画像から、YOLOv4 (Bochkovskiy et al. 2020) を用いて球果を検出するモデルを作成し、画像上にあった球果の 80~90% 程度を検出することができたことを報告している。しかし、単年度に取得した画像のみで学習と精度検証を行っており、作成したモデルが全く別の機会に撮影した画像に適用可能であるか、その汎用性を確認すべきことを指摘している。この確認は今後の実用化にあたって極めて重要な検討課題である。また、晴天日の撮影で白飛びしたシュートを誤検出した例や、密集して個々の球果の輪郭が不鮮明であった場合などには球果を検出できなかった例もあった。それゆえ、学習画像を追加することによる球果の検出精度の向上度合いを検討することなども、重要な課題として残されていた。

YOLOv4 は、C 言語をベースとしたニューラルネットワークのオープンソースフレームワークである darknet により実装されたモデルが公式に公開されており、花岡

(2021) ではそれを用いてモデルを作成した。darknet 版の YOLOv4 に関しては、特に OS が Windows の場合に実行環境の構築が非常に煩雑であるなど、作成したモデルを一般に普及する上では課題があった。一方で、YOLOv4 以降に公開された YOLOv5 (Jocher 2021) では、プログラム言語の Python のオープンソース機械学習ライブラリである PyTorch (Paszke et al. 2019) を用いて実装されたモデルが公式に公開されており、実行環境の構築が容易かつ汎用性が高い。Microsoft COCO dataset を用いた検証では、YOLOv4 の方が YOLOv5 よりも精度と速度の両面でわずかに優れていたという報告もあるものの (Nelson and Solawetz 2020)、同程度の精度が確保できるのであれば、YOLOv5 の利用はモデルの普及に対して大きな利点となり得る。また、樹冠にあるトドマツの球果のように、小さい物体が画像上にたくさん存在するような場合の物体検出に対して、YOLOv5 がどの程度の精度を示すのかを事例報告することは、育種や生態学分野の様々な研究等への利用可能性を検討する上でも有益な情報提供となるであろう。

本研究では、今後の YOLOv5 の利用可能性や実用性について検証することを主目的として、UAV を用いて撮影したトドマツの樹冠画像から球果を検出するモデルの作成を通して、(1)YOLOv5 を用いて得られた球果の検出精度を既存の YOLOv4 のモデル (花岡 2021) の精度と比較するとともに、(2) 新たに取得した画像を加えて学習した場合の球果の検出精度の向上度合いを検証した。また、YOLOv5 には、ニューラルネットワークの中間層のレイヤーサイズが異なる 5 種類のモデルが用意されており、レイヤーサイズが小さいものから順に YOLOv5n (213 レ

原稿受付：令和 4 年 4 月 13 日 原稿受理：令和 4 年 9 月 1 日

1) 森林総合研究所 林木育種センター 北海道育種場

2) 森林総合研究所 林木育種センター 遺伝資源部

\* 森林総合研究所 林木育種センター 北海道育種場 〒069-0836 江別市文京台緑町 561-1、E-mail: sohana@affrc.go.jp

イヤー)、YOLOv5s (213 レイヤー)、YOLOv5m(290 レイヤー)、YOLOv5l (367 レイヤー)、YOLOv5x (444 レイヤー)がある。そこで、(3)YOLOv5 についてはニューラルネットワークの中間層のレイヤーサイズが異なるモデルを用いた場合の精度差についても検証した。

## 2. 材料と方法

### 2-1. 学習・検証用データの作成と YOLOv5 のモデル作成

本研究では、花岡 (2021) で利用した、2020 年 6 月から 8 月に 12 ケ所のトドマツ採種園及び人工造林地で、DJI 社の Mavic2 Pro または Phantom 4 を用いて撮影し、作成した画像データ (以降、セット 1 とする) をそのまま利用するとともに、2021 年 6 月から 8 月にかけて、3 ケ所の採種園及び 3 ケ所の人工造林地 (Table 1) で撮影された画像 (以降、セット 2 とする) を新たに準備した。画像は、独自に撮影したものと、北海道森林管理局の協力を得て撮影・入手したものが含まれるが、いずれの場合もカメラを鉛直方向とし、可能な限り樹高から 5~10m 程度の高度で撮影すること、撮影時の露出設定はオートとするようにした。なお、画像撮影の依頼時の指示不足により、晴れの天候にも関わらず露出補正值がプラスとなっていた場合などもあり (Table 1)、樹冠の一部及び球果が白飛びした場合なども含まれたが、汎用性の高いモデルの作成に有用であると判断し、それらの画像も学習に利用した。撮影された球果は、UAV の直下で、画像の中心部にあった場合は主に真円状に写っており、また、UAV 直下から離れた位置で、画像の辺縁部にあった場合は、側面部が見えて楕円形に写っていた。なお、2021 年 (セット 2) の撮影では、画像上の球果の判読性を高めることを目的に、可能な限り高度を下げて撮影するよう、撮影協力者に依頼した結果、球果の短径 (約 3 cm) のピクセル

サイズが 40 ピクセル程度、画素寸法から換算して球果からの距離が 5m 未満であったと推察される画像も多数含まれていた。解析用画像の作成にあたっては、画像サイズを YOLOv4 のデフォルトサイズである 416 × 416 ピクセルとし、個々の撮影画像から球果が写っている部分を、画像の中心部と辺縁部の両方のパターンを含むように、また、画像の切り出し範囲が互いに重ならないように、画像あたり 2~6 ケ所を切り出して解析用画像とした。作成した画像 (Fig. 1) は、セット 1 で合計 448 枚、セット 2 で合計 174 枚となった。セット 2 の作成にあたっては、モデルの汎用性を高める観点から、2020 年の撮影画像にはあまり含まれていなかった、低空からの撮影で非常に大きく写った球果や、花岡 (2021) のモデルでは学習が不十分で検出の失敗例が多かった、密集した球果が写った箇所などを積極的に含むようにトリミングし、174 枚のうち 61 枚はこのような画像となっていた (Fig. 1 の下段)。

作成した全ての画像について、画像アノテーションソフトウェアである LabelImg (Tzutalin, LabelImg, Git code 2015) を用いて矩形のアノテーションデータを作成した。セット 1 の学習用画像については、356 枚の画像で合計 6138 個の球果数、検証用画像については 92 枚の画像で合計 1692 個の球果数となっていた (花岡 2021)。セット 2 については、目視で確認できた合計 2818 個の球果にアノテーションを付した上で、約 8 割にあたる 140 枚、合計 2308 個の球果が写った画像を無作為に抽出して学習用画像とし、残り 34 枚、合計 513 個の球果が写った画像を検証用画像とした。なお、大きく写った球果や密集した球果の数は、セット 2 の検証用画像に 13 枚含まれていた。

準備した学習用画像と検証用画像を、Table 2 の通り 13 通りの場合分けを行った上で、それぞれについて球果を検出する精度を検証した。YOLOv4 を用いた検証につ

Table 1. 撮影場所と撮影条件

撮影地	撮影機種	撮影日	露出補正值	天候	使用画像数
北海道育種場	Mavic2Pro	6/12/2020	0	晴れ	5
岐阜採種園	Phantom4	7/1/2020	0	晴れ	52
発足採種園	mavic2Pro	8/18/2020	0	晴れ	28
銀山採種園	Phantom4	8/21/2020	-0.7	晴れ	5
雨紛採種園	Mavic2Pro	7/9/2020	0	曇り	12
塩狩採種園	Mavic2Pro	8/11/2020	0	晴れ	2
乙部採種園	Mavic2Pro	7/15/2020	0	曇り	2
渡島森林管理署 2098 林班	Mavic2Pro	7/10/2020	0	晴れ	1
渡島森林管理署 5101 林班	Mavic2Pro	8/3/2020	0	晴れ	8
上川南部森林管理署 1065/1067 林班	Mavic2Pro	8/13/2020	0	曇り	27
上川南部森林管理署 1103 林班	Phantom4	8/21/2020	1	晴れ	1
檜山森林管理署 1409 林班	Phantom4	7/13/2020	-1.3	曇り	3
雨紛採種園	Phantom4	8/6/2021	0.3	晴れ	8
銀山採種園	Mavic2Pro	7/15/2021	0	晴れ	12
岐阜採種園	Phantom4	8/4/2022	0.3	晴れ	8
空知森林管理署 36 林班そ小班	Mavic2Pro	6/21/2021	-0.3	曇り	6
上川南部森林管理署 1231 林班は小班	Phantom4	8/27/2021	0	曇り	6
網走西部森林管理署 351 林班ち小班	Mavic2Pro	7/30/2021	-0.3	曇り	14

使用画像数は解析用画像の作成に使用された撮影画像枚数を示す。



いては花岡 (2021) で作成したモデルを活用し、YOLOv5 を用いたモデルについては本研究で新たに作成した。YOLOv5 の各モデルは開発者が公開する github (Jocher 2020) から入手し、利用した。YOLOv5n、YOLOv5s、YOLOv5m 及び YOLOv5l を用いた学習は、GPU として GeForce GTX1650 を搭載した PC を利用し、Python3.8.12 及び PyTorch1.7.1 を用いて実行した。GPU のメモリ容量の制限のため、学習時の batch size を YOLOv5n と YOLOv5s の場合には 16、YOLOv5m と YOLOv5l の場合には 8 とした。YOLOv5 における学習時の batch size は、初期の学習効率に影響するものの、学習の epoch 数を十分に大きくした場合には最終的な精度は変わらなかった

ことが報告されているため (Jocher 2021)、学習の epoch 数を同検証事例よりも大きい 300 に設定するとともに、画像サイズを 416 に変更して実行した。YOLOv5x の学習については、前述の PC では batch size を 8 としても GPU のメモリ容量が不足したため、GPU として GeForce RTX 3090 を搭載した異なる PC を利用し、Python3.8.12 と PyTorch1.10.1 を使用して、batch size を 32、epoch 数を 300、画像サイズを 416 に設定して実行した。

精度検証にあたっては、2つの矩形領域の重なりを示す指標である Intersection over Union (IoU) を基に、検出した矩形候補領域が正解の矩形領域と  $\text{IoU} \geq 0.5$  で重なっていた数を true positive (TP) とし、また、検出した矩形

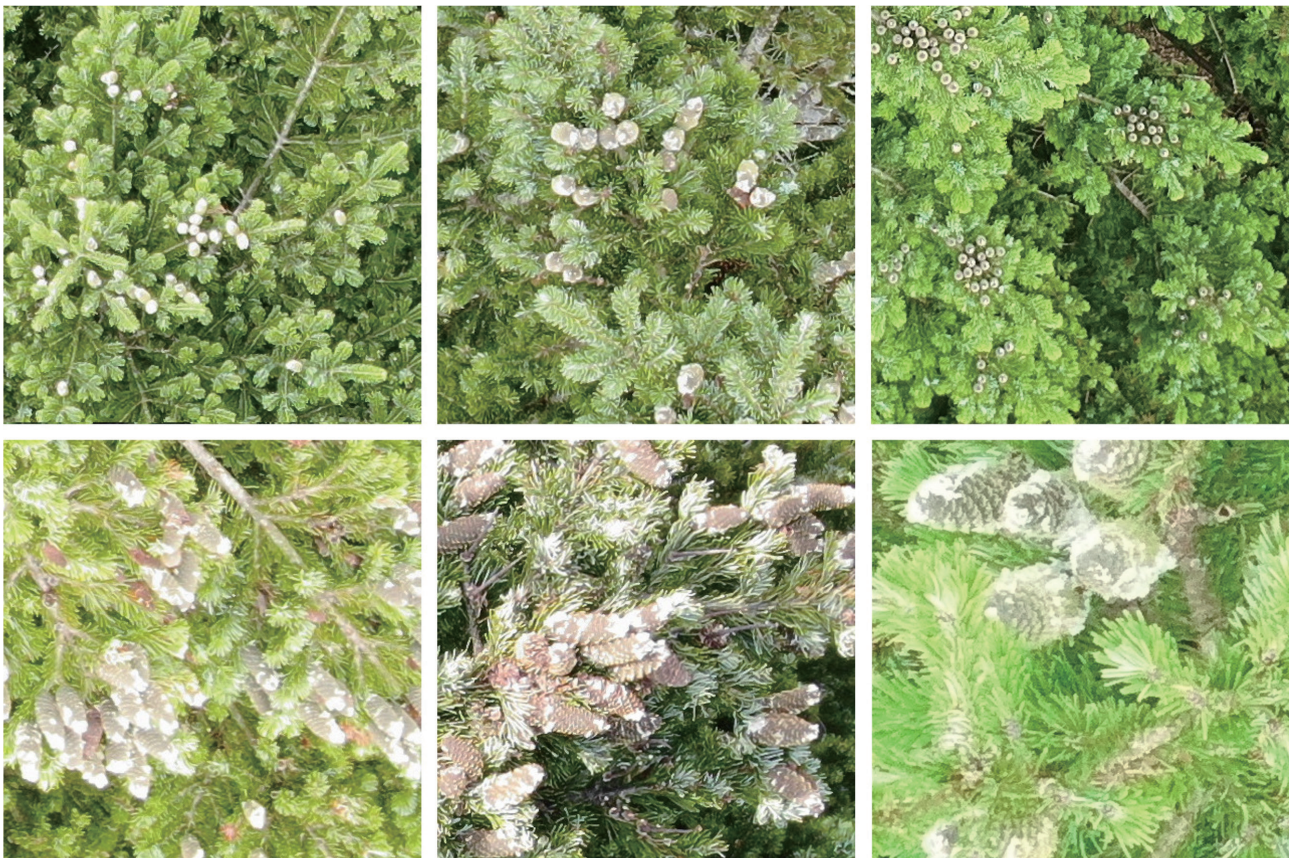


Fig. 1. 作成した解析用画像の例

Table 2. 作成した全モデルの一覧と得られた精度

検証	モデル	学習セット	検証セット	Precision	Recall	AP	備考
Case 1	YOLOv4	Set1	Set1	0.80	0.91	0.88	花岡 (2021) の結果
Case 2	YOLOv5m	Set1	Set1	0.86	0.84	0.90	
Case 3	YOLOv4	Set1	Set2	0.70	0.62	0.61	
Case 4	YOLOv5m	Set1	Set2	0.76	0.60	0.71	
Case 5	YOLOv4	Set1	Set1+2	0.78	0.84	0.82	異なる PC を利用
Case 6	YOLOv5m	Set1	Set1+2	0.81	0.81	0.85	
Case 7	YOLOv5m	Set1+2	Set1	0.85	0.89	0.93	
Case 8	YOLOv5n	Set1+2	Set1+2	0.84	0.84	0.89	
Case 9	YOLOv5s	Set1+2	Set1+2	0.87	0.83	0.90	
Case 10	YOLOv5m	Set1+2	Set1+2	0.85	0.85	0.90	
Case 11	YOLOv5l	Set1+2	Set1+2	0.85	0.85	0.90	
Case 12	YOLOv5x	Set1+2	Set1+2	0.86	0.85	0.90	
Case 13	YOLOv5m	Set2	Set1	0.80	0.71	0.77	

候補領域に対して正解の矩形領域が  $IoU < 0.5$  の重なりであったものを false positive (FP)、正解の矩形候補領域に対して検出した矩形候補領域が  $IoU < 0.5$  であったものを false negative (FN) として、各モデルを検証用画像に適用した場合の、それぞれの数を確認した。実質的には、FP は球果以外の物体を検出した数、FN は画像上に存在した球果を検出できなかった数となっていた。作成された各モデルの精度については、 $TP/(TP + FP)$  として計算される precision (P) と  $TP/(TP + FN)$  として検出すべき矩形領域のうち実際に矩形領域が検出された割合として計算される recall (R) 及び、両者を統合した指標である mean average precision (mAP; ただし、本研究では 1 クラスのみなので AP) を基に検証した。なお、これらの確認にあたっては、学習の epoch 毎に検証用画像で推論を行い、AP が最大となった epoch のベストモデルを用いて行った。

## 2-2. YOLOv4 と YOLOv5 の比較

本研究では、詳細は後述するが、結果として YOLOv5s から YOLOv5x には精度に差が出ない結果となったため、中間的なレイヤーサイズの YOLOv5m を用いて作成したモデルと花岡 (2021) で作成した YOLOv4 のモデルとを比較した。まず、YOLOv4 と YOLOv5m のそれぞれのモデルについて、花岡 (2021) の画像データセット (セット 1) のみを用いて学習・検証した場合の精度を比較した (Table 2 のケース 1 と 2)。また、セット 1 の学習用画像のみを用いて学習した YOLOv4 と YOLOv5m のモデルを、2021 年に撮影した検証用画像 (セット 2) に適用することで、それぞれのモデルの汎用性についても比較した (Table 2 のケース 3 と 4)。次に、セット 1 の学習用画像を基に作成した YOLOv4 と YOLOv5m のモデルを、セット 1 と 2 の両方を 1 つのセットとしてまとめた検証用画像に適用することで、検証用画像全体に対する精度差についても比較した (ケース 5 と 6)。その他、YOLOv4 と YOLOv5 のそれぞれが最大限の能力を発揮できる環境においては、YOLOv4 のほうがわずかに速度に優れるという報告があるものの (Nelson and Solawetz 2020)、実用的な利用への参考として、ケース 3 と 4 の 34 枚の検証用画像の推論を、GPU を搭載しない PC (MacBook Pro) で CPU のみ (Intel Core i5, 2.3GHz メモリ 16GB) で実行した際に要した時間を計測し、比較した。また、GPU による推論時間短縮効果を確認するため、GPU (GeForce GTX1650) を搭載した PC でもケース 4 の推論に要した時間を計測し、上述の CPU のみの PC の結果と比較した。

球果検出にあたっては、個々の矩形領域の正確さと各クラスの予測確率から算出される信頼度スコアの閾値を設定するが、ケース 1 ～ 4 については、花岡 (2021) との直接的な比較を行う観点から、球果検出時の信頼度スコアの閾値を YOLOv4 のデフォルト値である 0.25 とした。一方で、今回作成した YOLOv5 のモデルでは、信頼度スコアの閾値を 0.5 ～ 0.6 程度とした場合に P と R の

調和平均値 ( $F1$  スコア;  $(P \times R) / [(P + R)/2]$ ) が最大化され、両者のバランスが良くなると判断されたため (Fig. 2)、ケース 5～12 については、信頼度スコアの閾値を 0.5 に固定して検証した。

## 2-3. 学習用画像を増やした効果の検証

YOLOv5m を用いたモデル作成で学習用画像を増やした効果を検証するため、2020 年に撮影したセット 1 と 2021 年に撮影したセット 2 の両方の学習用画像を 1 つのセットとしてまとめて学習したモデルを、セット 1 の検証用画像のみに適用し (ケース 7)、ケース 2 の結果と比較した。同様に、セット 1 の学習用画像のみを用いて学習した YOLOv5m のモデルをセット 1 と 2 の両方を 1 つのセットとしてまとめた検証用画像に適用した場合 (ケース 6) と、セット 1 と 2 の両方を 1 つのセットとしてまとめた学習用画像を用いて学習したモデルをセット 1 と 2 の両方をまとめた検証用画像に適用した場合 (ケース 10) とを比較し、これらの結果を基に、画像を追加して学習させた場合の精度の向上について検討した。また、参考として、YOLOv5m でセット 2 の学習用画像のみを用いて学習したモデルを、セット 1 の検証用画像に適用した場合の精度も検証した (ケース 13)。

## 2-3. YOLOv5 のモデル比較

YOLOv5 の 5 つのモデル比較については、セット 1 と 2 の両方の学習用画像を用いて学習を行い、セット 1 と 2 の両方を含む検証用画像に適用し、それぞれから得られた精度を比較した (ケース 8 ～ 12)。

## 3. 結果と考察

### 3-1. YOLOv4 と YOLOv5m の比較

セット 1 の画像で学習と検証を行ったケース 1 (YOLOv4) とケース 2 (YOLOv5m) との比較では、P は YOLOv5m で 0.06 高くなる一方で、R は YOLOv4 が 0.07

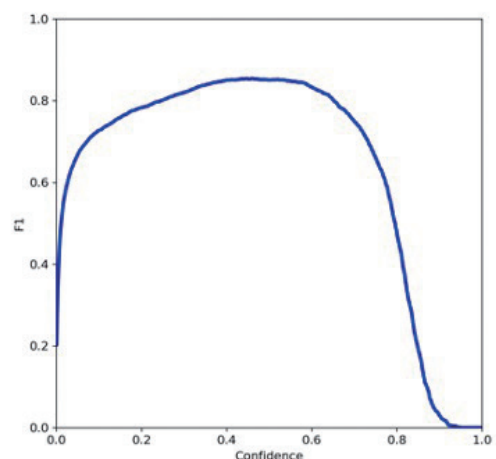


Fig. 2. ケース 10 における球果検出時の信頼度スコアの閾値と F1 スコアとの関係



高くなる結果となった。つまり、両者を比較した場合には、YOLOv5の方が球果以外の物体を誤検出するFPの割合が少ない一方で、球果を検出できなかったFNが多くなる傾向にあった。また、APはYOLOv4で0.88であったのに対し、YOLOv5mでは0.90となった。画像上に多数ある小型の物体（球果）を検出するモデルを作成した場合、YOLOv4とYOLOv5mの精度差はいずれの指標でも小さいものの、総合的に見てわずかにYOLOv5mの方が良い精度が得られた。しかし、どちらのモデルもPとRの両方が0.8以上かつAPも約0.9であり、両モデルとも実用的な利用が十分に可能なレベルにあると考えられた。また、実際に球果の付き具合を苗木生産者に情報提供する観点では、過大評価が抑制される方が好ましいとも考えられ、YOLOv5mの方が過大評価が生じにくいことは、実際の利用用途に適した特性と言えた。

ケース3と4について、GPU非搭載のPCで34枚の検証用画像の推論にかかった時間を比較したところ、YOLOv4では277秒、YOLOv5mでは27秒となり、YOLOv5mの方が大幅に短い時間で推論が完了した。CPUのみのPCでの実行など、実行環境によってはPyTorchで実装されたYOLOv5に優位性があると言えた。また、ケース4の検証用画像の推論をGPU (GeForce GTX1650)を搭載したPCで実行した場合には、推論に要した時間は6秒となり、推論時間がさらに短縮された。

### 3-2. 画像を追加した再学習の効果の検証

ケース1と3及びケース2と4を比較した場合には、ケース3と4でPとRの両方が0.1程度低くなっており、つまり、2020年に撮影されたセット1のみで学習したモデルを2021年に撮影されたセット2の検証用画像に適用し

た場合には、検出精度が10ポイント以上低い結果となった。また、セット2の学習用画像で学習させ、セット1の検証用画像に適用した場合のAPは0.77となり（ケース13）、ケース3と同様にAPが低い結果となった。これらの結果から、それぞれの年度に取得した画像だけで学習したモデルに十分な汎用性があったとは言えず、改善の余地があったことが示された。一方で、セット1とセット2の両方の学習用画像で学習したYOLOv5mのモデルをセット1の検証用画像に適用した場合には（ケース7）、Pが0.85、Rが0.89となり、ケース2と比較してRの値が高かった。結果として、APも0.93に向上した。これらの比較からは、画像を追加して行った学習が特にFNの減少に有効であったことが確認された。同様のことは、ケース6と10の比較からも言え、セット1と2の両方の学習用画像で学習したモデル（ケース10）の方が、セット1のみで学習したモデル（ケース6）よりもPとRの両方が4ポイント高かった。セット2の画像データは、花岡(2021)で課題とされていた密集した球果の画像を多数含むようにしたことに加え、セット1にはほとんど含まれていなかった低空からの撮影で大きく写った球果の画像なども多数含むようにしたことで、より汎用性の高いモデルに更新するために有効な追加の学習データであったと推察された。例として、ケース6のモデルでは検出できなかったが、ケース10のモデルでは検出できていた球果の例をFig. 3に示す。ケース10のモデルでは、密集した2~3個の球果がまとめて1クラスとして検出されている場合が散見されるものの、大幅に検出精度が向上していたことが見てとれた。既存のモデルを新たに撮影した球果画像に適用して検出精度が芳しくない場合であっても、それら新規に撮影した画像の一部などを用いて学

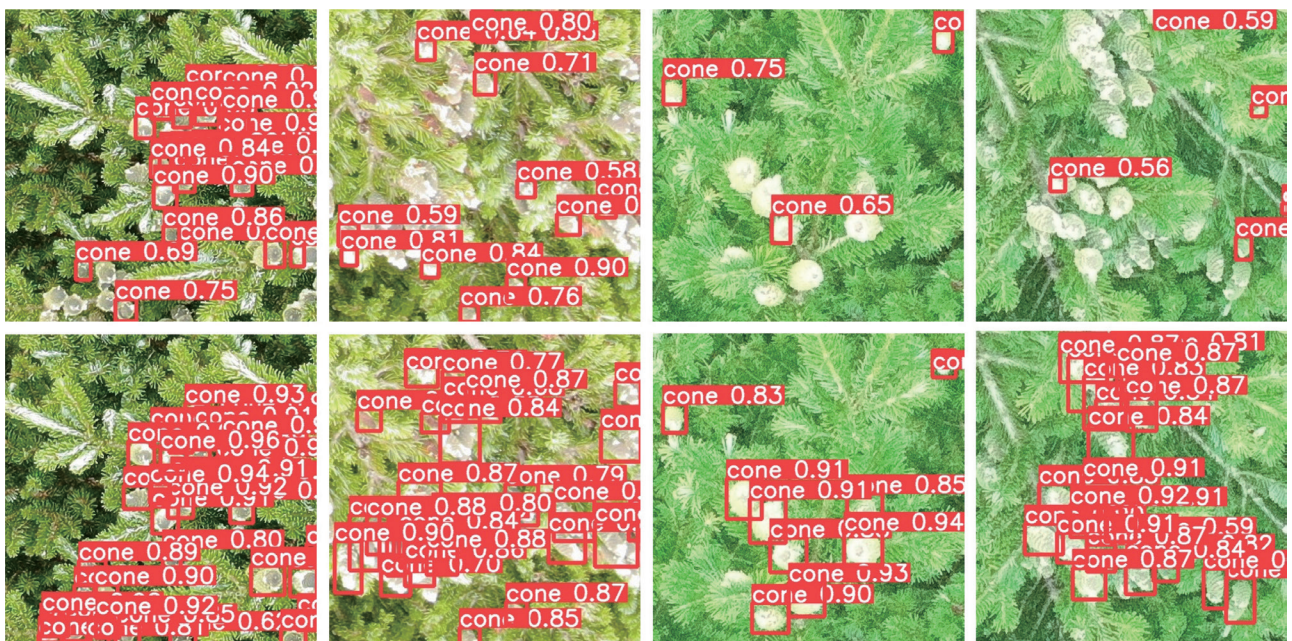


Fig. 3. ケース6のモデルによる検証用画像の推論結果の例（上段）とケース10のモデルによる推論結果の例（下段）。矩形で囲った球果それぞれに対して信頼度スコアが表示されている。





Fig. 4. ケース 10 のモデル (YOLOv5m) の検証用画像の推論結果の例。元画像を上段に、それぞれの推論結果を下段に示した。矩形で囲った球果それぞれに対して信頼度スコアが表示されている。

習をやり直すことで、AP を 0.9 程度にまで向上させ得ることを確認できた。一方で、上述のようにセット 2 による学習のみではセット 1 の AP は低かったことから (ケース 13)、広範な撮影条件に加えて、学習に用いる画像枚数も重要であることが示唆された。

一般に深層学習を含む機械学習においては、学習データの範囲外の予測は困難であることから (宮本 2020)、実用的に起こりうるトドマツの着花状況や撮影条件を含みつつより多くの画像を含むデータセットの構築が今後の精度向上の鍵であり、広範な撮影条件の追加と画像枚数の増加によりモデルをアップデートしていくなど、実装の中で精度向上を加えていくような取り組みが有効であると考えられる。

### 3-3. YOLOv5 のモデル比較

Microsoft COCO dataset など、多様で複雑なデータセットを取り扱った場合には、YOLOv5 のモデル間で mAP に大きな差があったことが示されており、レイヤーサイズの大きいモデルの方が精度が高くなるが、学習及び推論に要する時間も長い傾向にある (Jocher 2021)。本研究でセット 1 と 2 の全ての学習用及び検証用画像を用いて、YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x の比較を行った結果 (ケース 8 ~ 12) では、球果の検出精度に対するモデル間差は小さかった (Table 2)。最もレイヤーサイズが小さいモデルである YOLOv5n を用いた場合であっても、P と R の両方が 0.84、AP が 0.89 となった。YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x の比較では、P と R にわずかな違いが見られることがあったが、AP はいずれも 0.9 と差がなかった。例として、YOLOv5m を用いて様々な球果を検出した結果の例を Fig. 4 に示した。

トドマツの球果画像は、白飛びして写った球果とシュートや、樹冠下部にあった小さく見える球果とシュートなど、近くにあった確実に球果とわかるものと対比をしなければ目視でも見分けが難しい複雑な例が多数あり、検出誤差を生む主因となっている (花岡 2021)。今回の結果からは、これらの判別は少なくともレイヤーサイズの大きいモデルの使用によって解決することが難しいことが理解できた。今後の改善については、学習データを大幅に増やした場合の検討を進めることが必要と考えられる。

### 4. まとめ

本研究では、トドマツの樹冠画像から球果を検出することを通して YOLOv4 及び YOLOv5 の複数モデルの精度を検証し、全てのモデルで実用的なレベルの精度が得られた。また、GPU 非搭載の PC を用いた場合に YOLOv5 のモデルの方が YOLOv4 のモデルと比較して推論にかかる時間が短かったこと、また、YOLOv4 と比較すると OS によらず実装が容易であることなどもあり、実用面では YOLOv5 に優位性があると言えた。その他、YOLOv5n では球果の検出精度が低い結果ではあったものの、その差は他のモデルと比較して非常に小さかった。YOLOv5n のように軽量のモデルを用いることで、動画からのリアルタイム検出やタブレット端末等で動作可能なアプリケーションの作成なども実現できる可能性があり、このような発展的応用については今後の課題である。

### 謝辞

本研究に用いた画像データの収集にあたっては、北海道森林管理局の多大なるご協力をいただいた。また、本

研究の解析用画像の加工にあたっては、北海道育種場非常勤職員の田中昌子氏、中西裕子氏にご協力をいただくとともに、本稿執筆にあたって、北海道育種場育種課長の加藤一隆博士に助言をいただいた。ここに記し、感謝申し上げる。

#### 引用文献

- Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M. (2020) YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004. 10934.
- Chen, W., Zhang, J., Guo, B., Wei, Q. and Zhu, Z. (2021) An apple detection method based on des-Yolo v4 algorithm for harvesting robots in complex environment. *Mathematical Problems in Engineering*. ArticleID:7351470.
- Fan, Y., Zhang, S., Feng, K., Qian K., Wang, Y., Qin S. (2022) Strawberry maturity recognition algorithm combining dark channel enhancement and YOLOv5. *Sensors*, 22:419.
- 花岡 創 (2021) 畳み込みニューラルネットワークを用いた無人航空機 (UAV) 撮影画像からのトドマツ球果の検出. *日本森林学会誌*, 103, 372–377.
- Itakura, K., Narita, Y., Naoaki, S. and Hosoi, F. (2021) Automatic pear and apple detection by videos using deep learning and a Kalman filter. *OSA Continuum*, 4, 1688–1695.
- Jocher, G. (2021) “YOLOv5 Study: mAP vs Batch-size.”, <https://github.com/ultralytics/yolov5/discussions/2452>, (参照 2022-3-8).
- Liu, J. and Wang, X. (2020) Tomato disease and pests detection based on improved Yolo v3 convolutional neural network. *Frontiers in Plant Science*, 11, 898.
- 宮本 崇 (2020) パターン認識と法則発見のデータサイエンス. *AI・データサイエンス論文集*, J1, 270–277.
- Nelson, J. and Solawetz, J. (2020) “Responding to the controversy about YOLOv5.”, <https://blog.roboflow.com/yolov4-versus-yolov5/>, (参照 2022-3-7).
- Parico, A. I. B. and Ahamed, T. (2021) Real time pear fruit detection and counting using YOLOv4 models and deep SORT. *Sensors*, 21, 4803.
- Paszke, A., Gross, S., Massa F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Bai, J., Chintala, S. (2019) PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: *Advances in Neural Information Processing Systems 32* [Internet]. Curran Associates, Inc.; 2019. p. 8024–35.
- Tzutalin. LabelImg. Git code (2015). <https://github.com/tzutalin/labelImg>, (参照 2022-3-8).

## Model comparison of object detection algorithm YOLO based on the cone detection from the imagery of tree crowns of *Abies sachalinensis*

So HANAOKA<sup>1)\*</sup> and Eitaro FUKATSU<sup>2)</sup>

### Abstract

Differences in the accuracy of cone detection on the tree crown of *Abies sachalinensis* were compared using multiple models of the object detection algorithm “You Only Look Once” (YOLO). Training and validation were conducted using YOLOv4 and YOLOv5m, which both showed an average precision (AP) of around 0.9. Five models of YOLOv5 (YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x) with different neural network layer sizes were also compared. AP was 0.9 in all models except for YOLOv5n, and the accuracy was similar regardless of the neural network layer size.

**Key words :** YOLOv4, YOLOv5, object detection, deep learning, fruit set evaluation

---

Received 13 April 2022, Accepted 1 September 2022

1) Hokkaido Regional Breeding Office, Forest Tree Breeding Center (FTBC), Forestry and Forest Products Research Institute (FFPRI)

2) Genetic Resources Department, FTBC, FFPRI

\* Hokkaido Regional Breeding Office, FTBC, FFPRI, 561-1 Bunkyo-daimidori-machi, Ebetsu, Hokkaido, 069-0836 JAPAN  
; E-mail: sohana@ffpri.affrc.go.jp