H-025

複数の深層学習モデルを用いた広角画像からのブドウ葉検出手法の検討
A Study on Grape Leaf Detection from Wide-Angle Images Using Multi Deep Learning Models

三浦 怜† 小嶋 和徳† 伊藤 慶明†
Rei Miura Kazunori Kojima Yoshiaki Ito

1. はじめに

近年，農業従事者の高齢化や減少により，人手不足が深刻化している．この問題を解決するため，ドローンの活用が期待されているが，果樹園などの山間部圃場では導入が進んでいない．先行研究[1][2]では，ブドウ葉を対象にドローンで撮影した広角画像から手動で切り取ったブドウ葉画像に対し，複数の異常検知手法を組み合わせて異常葉を検出する手法の提案を行った．しかしこれらの研究では，広角画像からブドウ葉を自動で検出することは行っていない．本研究では，複数の物体検出手法と半教師あり学習を組み合わせ，ブドウ葉検出精度の向上を目指す．

2. 先行研究

2.1 複数の異常検知手法と U-Net を組み合わせたブドウ葉の健康状態識別手法[2]

先行研究[2]では切り取ったブドウ葉画像に対し，PaDiM，PatchCore，Cs-Flow と U-Net を組み合わせたブドウ葉の健康状態識別手法を提案した．複数の異常検知手法のヒートマップと U-Net の予測マスクを結合することで，画像の異常個所を強調し，異常識別を行った．この手法を用いることで，最大で Accuracy 0.94，Precision 0.96，Recall 0.97，F1 0.94 という結果が得られた．

2.2 ブドウ葉検出における現状の課題

ブドウ葉検出における現状の課題は，学習用画像が不足していることである．YOLO の公式ドキュメント[4]によると，1 クラス当たり 1,500 枚以上の学習用画像を使用することが推奨されている．しかし，本研究では学習用画像が 149 枚しか確保できていない．さらに，アノテーション作業には多くのコストがかかるため，画像数の増加には限界がある．そのため，画像数を増やすための効率的な方法を模索する必要がある．また，調査した限りでは，広角画像からのブドウ葉検出に関する研究は行われておらず，手法は未だ十分に検討されていない．

3. 提案手法

図 1 に提案手法の流れを示す．本研究では，複数の物体検出手法を用いた半教師あり学習により，学習画像を拡張することでブドウ葉検出精度の向上を図る．用いる画像は，先行研究と同様に，動画から切り出した広角画像である．まず，1. 複数の物体検出モデルを，学習用画像で学習する．次に，2. ラベルがついていない広角画像に対して疑似的にラベルを付与し，学習用画像を拡張する．疑似ラベルの付与には，学習済みモデルの推論結果を Weighted Boxes Fusion（WBF）[5]で統合した結果を用いる．次に，3. 学習用画像と疑似ラベル付き画像を用いて，学習済みモデルを再学習する．最後に，4. 評価用画像により，再学習済みの物体検出モデルの精度を評価する．

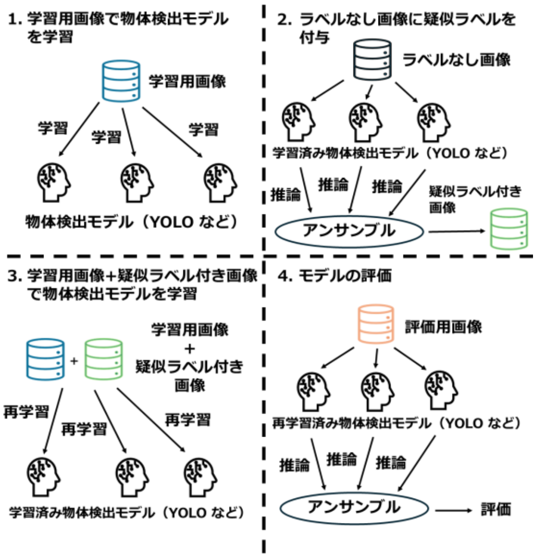

図 1　提案手法概要図

3.1 WBF(Weighted Boxes Fusion)[5]

WBF は，複数のモデル出力を統合し，精度を向上させるための後処理手法である．図 2 に Non-Maximum Suppression（NMS）と WBF の処理の流れを示す．NMS では，最も信頼度スコアの高いバウンディングボックスを残し，他の重複したバウンディングボックスを削除する．一方，WBF では同じオブジェクトに対応するバウンディングボックスを重み付きで統合する．各バウンディングボックスの重みは，検出信頼度スコアやモデルごとの重みを考慮して重みづけを行う．これにより，複数のモデルからの出力を柔軟に反映できる．特に，異なるモデル間の相補的な情報を活用できるため，アンサンブル学習において有効である．

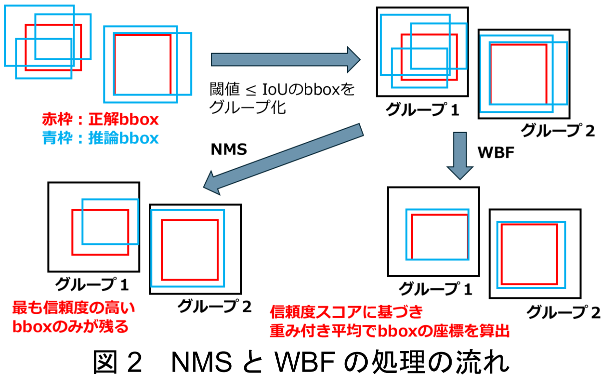

図 2　NMS と WBF の処理の流れ

†岩手県立大学　Iwate Prefectural University

Copyright © 2025 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

4. 実験

4.1 評価方法

データセットは 187 枚のラベル付き画像を学習用画像 149 枚，検証用画像 19 枚，評価用画像 19 枚に分割した．学習用画像にはデータ拡張を施し，8 倍の 1192 枚に拡張した．データ拡張には，切り抜き，水平反転，色調整を使用した．また，疑似ラベルを付与する画像として，新たに動画から 1782 枚の画像を切り出した．物体検出モデルとして，YOLOv8[3]，Faster R-CNN[6]，EfficientDet[7]を使用した．各モデルの学習後，評価用画像を用いて F1 スコア，mAP@50，mAP@50:95 を算出した．また，WBF を適用し，推論結果を統合したアンサンブルモデルの評価も行った．

4.2 ラベル付き画像による物体検出モデルの学習

ラベル付きの学習用画像のみを用いて，物体検出モデルの学習を行った．評価画像での評価結果を表 1 に示す．最高精度は赤字で示している．

表 1　ラベル付き画像で学習したモデルの評価

	F1	mAP@50	mAP@50:95
YOLOv8	0.822	0.863	0.622
Faster R-CNN	0.809	0.843	0.585
EfficientDet	0.774	0.838	0.573
アンサンブル	**0.838**	**0.896**	**0.639**

4.3 YOLOv8 を用いた半教師あり学習

ラベル付きの学習用画像と疑似ラベル付き画像を用いて，4.2 の学習済み物体検出モデルを再学習した．疑似ラベルの付与には，4.2 で学習済みの YOLOv8 を使用した．評価画像での評価結果を表 2 に示す．最高精度は赤字で示している．

表 2　YOLOv8 で付与した
疑似ラベルで再学習したモデルの評価

	F1	mAP@50	mAP@50:95
YOLOv8	0.826	0.880	0.647
Faster R-CNN	0.836	0.875	0.626
EfficientDet	0.817	0.882	0.608
アンサンブル	**0.844**	**0.909**	**0.654**

4.4 複数の物体検出手法を用いた半教師あり学習

疑似ラベルの付与に 4.2 の学習済みのアンサンブルモデルを使用し，4.3 と同様の実験を行った．評価画像での評価結果を表 3 に示す．最高精度は赤字で示している．

表 3　アンサンブルモデルで付与した
疑似ラベルで再学習したモデルの評価

	F1	mAP@50	mAP@50:95
YOLOv8	0.844	0.908	**0.665**
Faster R-CNN	0.830	0.884	0.621
EfficientDet	0.790	0.865	0.576
アンサンブル	**0.852**	**0.913**	0.645

4.5 考察

表 1，表 2，表 3 の結果から，半教師あり学習を用いた手法は，ラベル付き画像のみの学習と比較して全評価指標で高精度を示した．YOLOv8 では，複数の物体検出手法による半教師あり学習の方が，YOLOv8 単体による半教師あり学習よりも全評価指標で高精度を示した．一方で，Faster R-CNN と EfficientDet の一部のケースでは，複数の物体検出手法による半教師あり学習よりも YOLOv8 を用いた半教師あり学習の方が高精度を示した．この結果から，疑似ラベル生成に用いるモデルの選定が物体検出精度に影響を与えると考えられる．さらに，追加の実験を通じて，疑似ラベル付与時の信頼度閾値が物体検出精度に影響を与えることを確認した．信頼度スコア 0.25 以上 0.75 未満の範囲では精度に大きな変化は見られなかったが，0.75 以上では精度が低下した．このことから，低〜中程度の信頼度にも有用な情報が含まれており，閾値設定を高くした場合に有効な疑似ラベルを除外する可能性があると考えられる．今後は，信頼性と情報量のバランスを考慮した疑似ラベルの選別方法の検討が必要であると考えている．

5. おわりに

本研究では，複数の物体検出手法と半教師あり学習を組み合わせることで，ブドウ葉検出精度の向上を図った．特に，疑似ラベルを活用することで，学習用画像の効率的な拡張および精度向上の実現が可能であることを確認した．さらに，疑似ラベル生成に用いるモデルの選定や，疑似ラベルの信頼性が物体検出性能に与える影響を確認した．今後は，疑似ラベル生成に用いる適切なモデルの選定や信頼性の高い疑似ラベルの生成手法について検討する．

参考文献

[1] 下總有芽, 小嶋和徳, 伊藤慶明: PaDiM を用いたブドウ葉の健康状態識別システムの提案, FIT2024 第 23 回情報科学技術フォーラム, 第 3 分冊, H-058, pp.241--242, 2024-9.
[2] 下總有芽, 小嶋和徳, 伊藤慶明: 複数の異常検知手法と U-Net を組み合わせたブドウ葉の健康状態識別手法, 情報処理学会第 87 回全国大会, 2Y-01, pp.2-1031--1032, 2025-3.
[3] Jocher, G., Chaurasia, A., Qiu, J.:Ultralytics YOLOv8, version 8.0.0, available from ＜ https://github.com/ ultralytics/ultralytics ＞, AGPL-3.0, (2023).
[4] Ultralytics:Tips for Best Training Results, Ultralytics YOLO Docs, Ultralytics YOLO Docs(online), available from ＜ https://docs.ultralytics.com/yolov5/tutorials/tips_for_best_training_results/ ＞ (accessed 2025-06- 02).
[5] Solovyev, R., Wang, W. and Gabruseva, T.:Weighted boxes fusion: Ensembling boxes from different object detection models, arXiv preprint arXiv:1910.13302, (2019).
[6] Ren, S., He, K., Girshick, R., et al.:Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, Proc.Advances in Neural Information Processing Systems 28 (NIPS 2015), pp. 91-99, (2015).
[7] Tan, M., Pang, R., Le, V.Q.:EfficientDet: Scalable and Efficient Object Detection, Proc.IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 10781-10790, (2020).

Copyright © 2025 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.