

• Single Shot MultiBox Detector を用いた高精度なトマトの花検出

中尾大輝¹ 峰野博史^{1,2}

¹静岡大学情報学部 ²JST さきがけ

1.はじめに

高品質なトマト栽培において、開花は生育ステージが大幅に変化する1つの要因であり、開花日は人工授粉のタイミングの判別や収穫日の予測において重要なデータとなる。しかし、現在の人手による開花日の記録には2つの課題がある。1つ目は、開花日の記録は人間が目視で行う必要があり、大規模な栽培を行う際には多大な労力を要することである。2つ目は、人間が目視で記録を行うので人為的な間違いが起こりうるため正確性に欠けることである。これらの課題は開花日を自動で記録することで解決できる。この開花日の自動記録を行うには、花を検出することが不可欠である。そこで本研究では、栽培中のトマトを撮影した画像に深層学習ベースの一般物体検出アルゴリズムであるSSD(Single Shot MultiBox Detector)[1]を適用することで花の検出を行う。SSDを用いることで高精度かつ高速な花の検出を行うことができる。

2.関連研究

現在、画像に含まれる花を検出する様々な手法が提案されている。トマトの花の検出に色情報を用いる手法[2]では、RGB画像をHSV画像に変換した後に黄色い部分を閾値で抽出して花を検出している。しかし、この手法は明るさの変化や画像内に花と同色の物体が映っていた場合に誤検出が発生する。一方、Sliding windowを用いて物体を検出する手法[3]では、様々なサイズの四角い領域を一定ピクセル毎に画像全体にスライドさせ、四角い領域内に対象物が存在するかをSVM等を用いて判別を行う。この手法では物体を検出する計算コストが大きく、検出できる形状やサイズに制限があるという課題がある。また、Selective searchを用いて物体を検出する手法[4]では、初めに画像を複数の小領域に分割し、各小領域のヒストグラムやテクスチャ等の類似度に応じて結合を繰り返して物体を検出する。この手法では、画像が高解像度のときに計算コストが大きく、低解像度のときに細かい物体検出ができなくなるという課題がある。以上から、一旦モデルを構築してしまえば、明るさの変化に影響を受けず、少ない計算コストで物体を検出することができる深層学習を用いた物体検出の適用を考える。

3.提案手法

明るさの変化に影響を受けず、低解像度の画像に対して高速かつ高精度に物体を検出できるSSD(Single Shot Multibox Detector)を用いてトマトの花の検出を行う手法を提案する。SSDは深層学習ベースの物体検出アルゴリ

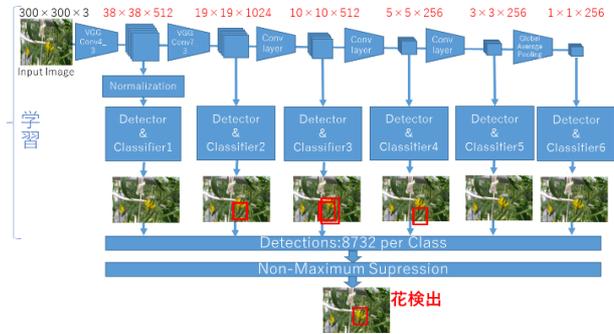


図1: SSDのアーキテクチャ

ズムの一種であり、物体候補領域の生成と分類を同時に学習・実行することが可能である。物体候補領域の生成と分類をend-to-endで学習することで高精度な検出を実現している。また、ニューラルネットワークのみで検出を行うため、GPUを用いれば高速に検出できる。

SSDのアーキテクチャを図1に示す。SSDの入力には300×300の画像を用いることとした。SSDでは、まず最初にImageNetで学習済みの既存モデルであるVGG[5]を用いて画像から特徴マップを生成する。次に、VGGで生成された特徴マップにCNNを多段的に適用しマルチスケールの特徴マップ(図1の上部四角部分)を生成する。最後に、上記で生成された各スケールの特徴マップにCNNを適用して物体の検出と分類を同時に行い、その結果を結合して最終的な検出を行う。SSDの出力結果には検出した花の座標とクラス信頼度が出力される。クラス信頼度は0から1の値を取り、値が1に近づくほど検出した物体が花である確率が高いと予測している。各スケールの特徴マップに適用するCNNのカーネルサイズは一定であるため、サイズの大きい特徴マップでは小さい物体を、サイズの小さい特徴マップでは大きい物体を検出する役割がある。

4.基礎評価

提案手法の評価として、トマトの花の座標とクラス信頼度を推定し、花検出精度をPrecision, Recallで評価した。評価の目的は学習データ拡張の効果と花のサイズ毎の精度比較である。ディープラーニングでは、学習に用いるデータ数が多いほど性能を発揮するため、データ拡張を適用する。データ拡張時のデータセットの詳細を表1に示す。データ拡張には、一般的に有効であると述べ

表1: データセットの詳細

データ種類	撮影場所	品種	データ数		
			拡張前	左右反転	左右反転+Random Erasing
学習データ	静岡大学農学部	cf桃太郎ヨーク	960件	1920件	2880件
検証データ	静岡大学農学部	cf桃太郎ヨーク	240件	480件	720件
評価データ	静岡県農林技研	桃太郎ピース	509件	509件	509件

A study on Single Shot MultiBox Detector for tomato flower detection

Taiki Nakao¹, Hiroshi Mineno^{1,2}

¹ Faculty of Informatics, Shizuoka University

² JST PRESTO

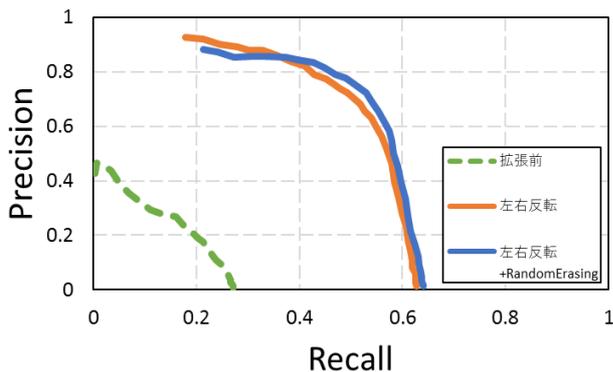


図 2：データ拡張の評価結果

表 2：花サイズのパラメータ

設定	花サイズ	データ数
All	All size	1228
Small	900(1%)pixels以下	299
Large	900pixels以上	929

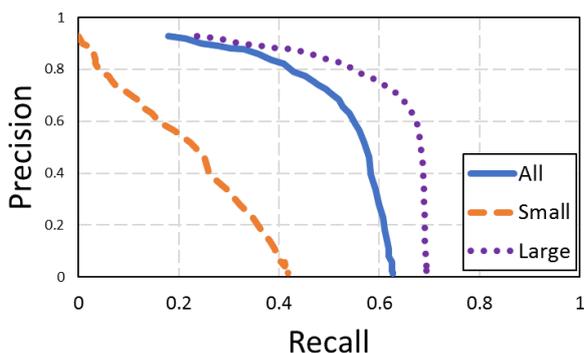


図 3：花サイズ毎の評価結果

られている画像の左右反転と、RandomErasing[6]を用いた。RandomErasing は画像中の物体検出対象物の一部をマスクする手法であり、対象物の一部が隠れている場合の検出に有効である。GoPro HERO4 を用いて、静岡大学農学部施設園芸環境で cf 桃太郎ヨークの花を含む画像を学習データ用に 960 件、検証データ用に 240 件収集した。また、静岡県農林技術研究所で桃太郎ピースの花を含む画像を評価データ用に 509 件収集した。目視では品種による花の形状や色の違いは見られなかった。撮影した画像のサイズは 2720×2040 である。各画像に人手で花の位置とクラスをラベル付けしてデータセットの作成を行った。

学習データと検証データの画像に対して、データ数を拡張した時の評価結果を図 2 に示す。図 2 から、データ拡張を行うことで花検出精度が向上することが確認された。RandomErasing を用いることで、Recall が 0.51 の時に Precision が 0.05 向上した。大量のデータを収集してラベル付けを行うのは困難であるため、Random Crop など多様なデータ拡張手法を適用してデータ数を増やすことで、さらなる花検出精度の向上が期待できる。

次に、画像に対する花のサイズによる検出精度の違いを比較する評価を行った。学習には左右反転で拡張したデータセットを用いた。比較したサイズのパラメータを表 2 に、評価結果を図 3 に示す。図 3 の結果からサイズ



図 4：花検出結果

の小さい花の検出が困難であることが確認された。小さい花の検出精度が低い理由として、SSD 入力時に画像を圧縮しているため、小さい花は特徴量が消滅しているからであると考えられる。入力画像サイズを上げる等の改善を検討する。

評価データに対する SSD の花検出結果例を図 4 に示す。花が存在すると判定された領域は赤いバウンディングボックスで出力される。

5. おわりに

本研究では、深層学習ベースの物体検出アルゴリズムである SSD を用いたトマトの花検出手法を提案した。基礎評価の結果、学習データを増やすことで精度が向上することが確認されたため、多様なデータ拡張手法を適用することで花検出精度の向上が期待できる。また、複数の DNN やアルゴリズムを組み合わせたアーキテクチャ [7] の構築を行い、花検出精度の向上を目指す。今後は、開花日自動記録システムへの拡張を考慮して、サイズの小さい花を検出する必要性と精度改善の検討を進める。花検出精度の向上を終えた後、トマトの花検出モデルを開花日自動記録システムへの適用を行う予定である。

謝辞

本研究の一部は JST さきがけの支援を受け実施された。

6. 参考文献

- [1] Wei Liu, et al., “SSD: Single Shot MultiBox Detector”, Computer Vision - ECCV 2016 pp 21-37, 2016.
- [2] Dor Oppenheim, et al., “Detecting Tomato Flowers in Greenhouses Using Computer Vision”, WASET International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:11, No:1, 2017.
- [3] Wei Guo, et al., “Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images”, BioMed Central, 2015.
- [4] Ross Girshick, et al., “Rich feature hierarchies for accurate object detection and semantic segmentation”, CVPR 2014, 2014.
- [5] Karen Simonyan, et al. “very deep convolutional networks for large-scale image recognition”, arXiv:1409.1556v6, 2014.
- [6] Zhun Zhong, et al., “Random Erasing Data Augmentation”, arXiv:1708.04896, 2017.
- [7] Xianzhi Du, et al., “Fused DNN: A deep neural network fusion approach to fast and robust pedestrian detection”, IEEE Winter Conf. on Applications of Computer Vision, 2017.