

ディープラーニングを用いた
エダマメ外観精選別用物体検出 AI の
学習データに関する研究

Studies on training data of green soybean appearance quality sorting
AI using deep learning

岩手大学大学院連合農学研究科
地域環境創生学専攻
地域環境工学連合講座

森 智洋

目次

第 I 章 序言	1
第 1 節 エダマメの歴史と生産体系について	1
(1) エダマメの歴史について	1
(2) 海外のエダマメ生産について	1
(3) 日本のエダマメ生産について	2
第 2 節 エダマメの栽培工程と課題	4
(1) エダマメの栽培工程と機械化について	4
(2) エダマメの出荷調製について	5
(3) 外観精選別作業の課題について	5
第 3 節 エダマメの精選別に関する研究	7
第 4 節 ディープラーニングと AI	9
(1) 第 3 次 AI ブームと CNN の基本構造	9
(2) 画像認識分野における AI の発展	12
(3) DL で開発した AI の応用事例	14
第 5 節 DL で開発した物体検出 AI の課題	15
(1) DL と従来の画像処理の違いについて	15
(2) DL における学習データの研究について	18
第 6 節 研究の目的と目標	21
第 7 節 論文の構成	22
第 II 章 学習データセットに含まれるエダマメ品種の違いが物体検出 AI の精 選別精度に与える影響	23
第 1 節 緒言	23
(1) 学習データの内容が AI の精度に影響した例について	23
(2) ネットワーク構造の異なる物体検出 AI	25

(3)本実験の目標	26
第2節 実験方法	27
(1) 供試材料と画像の収集	28
1) 本実験で使用したエダマメ品種と外観品質分類項目について	28
2) 本実験で収集した画像について	29
(2)アノテーションによる教師データの作成	31
(3) 品種の組み合わせを変えた7つのデータセットの構成	34
(4) DLによるAIの開発	35
(5) AIの精度評価	39
1) 適合率, 再現率, F値	39
2) ニュートン効率	40
3) 誤分類率と未検出率	41
4) 精度評価全般の補足について	42
第3節 結果と考察	44
(1) 適合率・再現率・F値	44
(2) ニュートン効率	50
(3) 誤分類率と未検出率	52
1) 誤分類率	53
2) 未検出率	56
3) AIのネットワーク構造の違いによる誤分類と未検出の内容の変化 について	57
第4節 摘要	58
第III章 CNNの分類根拠可視化技術を利用したAIの精度変化要因の解明 ...	59
第1節 緒言	59
(1) 第II章の実験結果とAIの精度変化の要因	59

(2) 説明可能な AI による分類根拠の特定	59
(3) 本実験の目標	62
第 2 節 実験方法	63
(1) 供試材料と画像の収集	65
(2) 学習データと精度評価用のデータの準備	69
(3) DL による AI の開発	70
(4) AI の精度評価	72
(5) Score-CAM による分類根拠の特定方法について	73
第 3 節 結果と考察	76
(1) AI の精度評価	76
(2) Score-CAM による各 AI の分類根拠の特定	84
1) 品種が分類根拠の変化に与える影響について	84
2) 品種バイアスの解消方法について	86
3) 学習データに潜む精度低下因子の特定手法について	88
第 4 節 摘要	90
第 IV 章 総合考察	91
第 1 節 エダマメの精選別用物体検出 AI の MLOps の考案	91
(1) 本研究で得られた知見	91
(2) エダマメ精選別用 AI の MLOps	92
第 2 節 DL で開発した物体検出 AI とオープンソースソフトウェアを搭載し た精選別システム	98
第 3 節 学習データの内容・構成に関する研究の促進	103
第 4 節 選別情報の二次利用	105
(1) 生産者と運送業者の各ニーズを満たす集荷システム	105
(2) 日ごとの収穫量の調整	107

(3) 選別項目の自動設定	107
(4) 栽培へのフィードバック	111
(5) 育種への応用	112
(6) 精選別以外の AI との連携.....	112
謝辞	115
引用文献	116

第 I 章 序言

第 1 節 エダマメの歴史と生産体系について

(1) エダマメの歴史について

エダマメ（学名：*Glycine max*）はマメ科の 1 年草であり，未成熟なダイズを収穫したものである¹⁾。ダイズの起源は紀元前 2000 年頃の中国にあり，日本には縄文時代の後期から弥生時代の前期にかけて中国より伝来した。日本でエダマメを食べるようになった時期は，江戸時代とされている。江戸時代は都市圏で外食文化が広まった時期であり，エダマメは主に露天で売られていた。当時のエダマメは，枝についたエダマメをそのまま茹でて，ファーストフードのようにして食べられていた。次第に，エダマメは一般的な食べ物として定着していった。海外に関して，エダマメは中国で紀元前 200 年頃に食用や薬用として食べられるようになった²⁾。その後，エダマメは華僑が東南アジアや中央アジア，そして欧米へ伝え，世界に広まった。近年，エダマメは日本食ブームの影響もあり，海外でも人気の食材となっている。エダマメの栄養価は高く，タンパク質や糖質，ビタミン B 群，ビタミン C をバランス良く含んでおり，さらに風味も優れている³⁾。エダマメは海外で日本食レストランなどを通じて広められ，「EDAMAME」として認知されつつある⁴⁾。世界のエダマメの消費量は，こうした認知度や人気の高まりによって，今後も増え続けると予想されている。

(2) 海外のエダマメ生産について

世界のエダマメ主要産地は東アジアや東南アジアに多く，日本や台湾，タイ，中国などがある。エダマメは世界で認知されているが，海外の消費量は日本よりも少ない。そのため，海外のエダマメは，冷凍して日本やアメリカなどへ輸出される割合が高い。台湾とタイのエダマメ生

産量はそれぞれ約 70,000t と 20,000t で、そのうち約 85~90%が冷凍で輸出される⁵⁾⁶⁾。このような背景から、海外のエダマメ生産者は、冷凍食品企業や輸出企業と契約しており、企業から出荷量の設定や生産補助を受けて栽培を行う。地域によっては、少数のリーダーが多数の生産者をまとめ、企業と契約を結んでいる。これによって、海外のエダマメ生産は大規模化が進み、生産者と企業が協力してエダマメのサプライチェーンを構築している。

(3) 日本のエダマメ生産について

日本は、エダマメの消費量が最も多い国の1つとされる⁷⁾。日本のエダマメ生産は、家族経営などによる小規模栽培が多い。これは、多くのエダマメが生鮮食品として国内流通されるためである。一方、海外は主に輸出のためにエダマメを栽培しており、生産者と企業が協力して大規模栽培を行っている。しかし、日本はエダマメを冷凍せずに市場へ流通させる割合が高く、生産者が個々に流通業者へ出荷している。以上から、エダマメの栽培目的や栽培規模は、日本と海外で異なっている。

日本のエダマメの栽培は、全国作付面積が 12,800 ha、全国出荷量が 51,200t で (2020 年)⁸⁾、近年は横ばいで推移している。国内のエダマメの主要産地は、新潟県や山形県、秋田県など水稻の生産が盛んな地域や、大都市周辺の地域に多い (図 1-1~1-2)。日本では 1940 年代後半の食糧難をきっかけに、コメの増産が推進された。しかし、1960 年代にコメの供給が需要を上回り、コメ余りが発生した。そこで、水稻が盛んであった地域は、コメの生産調整による水田転換の必要から、転換作物の導入を行った。当初の転換作物は、ムギやダイズ、そしてエダマメであった。1990 年代までの主な転換作物はムギやダイズであったが、これらは販売収益が低いことから次第に栽培が減少した。そこで、エダマメの

販売収益の高さが注目された⁹⁾。平成30年度のダイズの販売価格は148円/kgであるのに対し、エダマメの販売価格は691円/kgと約4倍である¹⁰⁾。また、エダマメは未熟のダイズであるため、播種作業や防除作業にダイズ用の機械を利用することができる。そのため、ダイズを栽培してきた生産者は、エダマメの導入が比較的容易である。

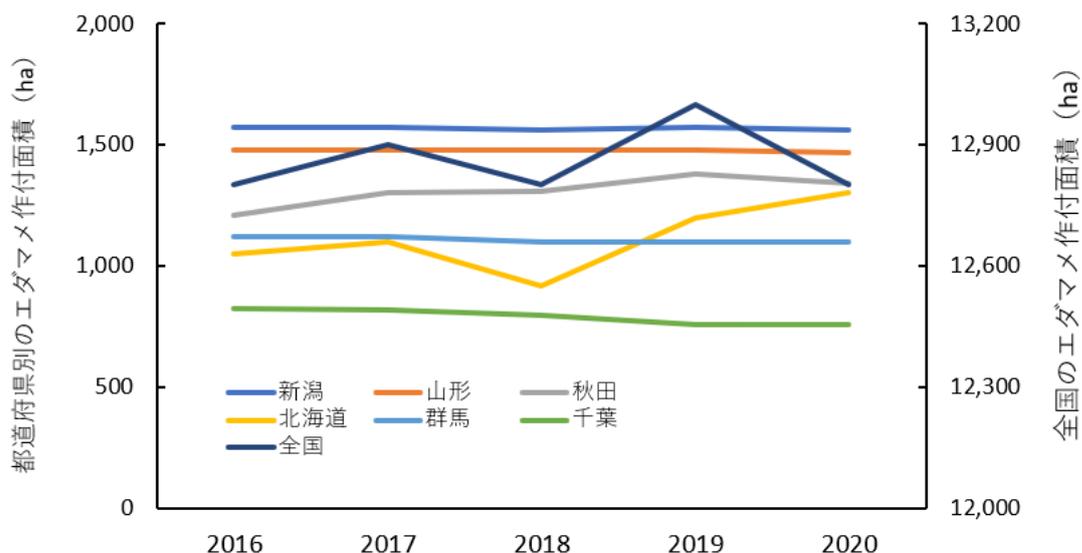


図 1-1 全国および各都道府県のエダマメ作付面積

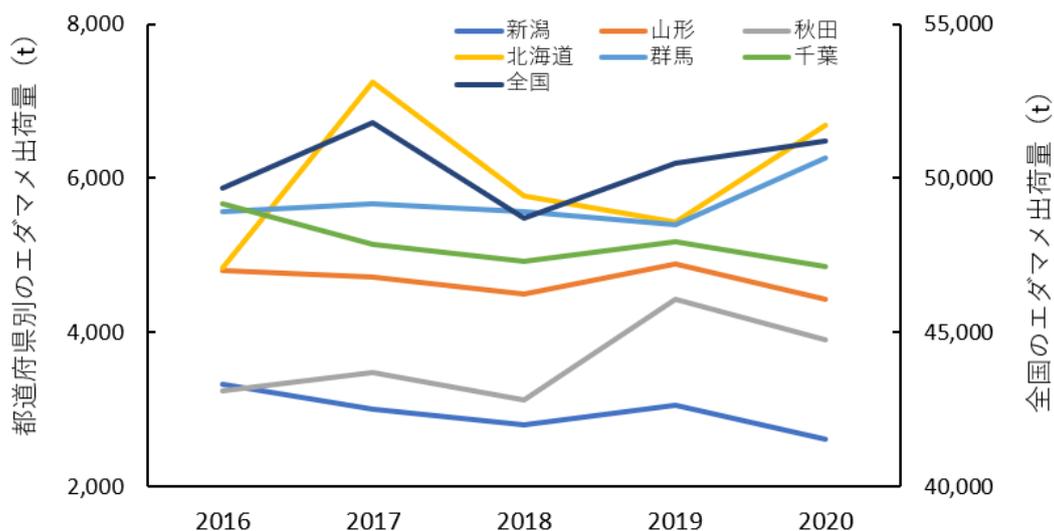


図 1-2 全国および各都道府県のエダマメ出荷量

第2節 エダマメの栽培工程と課題

(1) エダマメの栽培工程と機械化について

エダマメの栽培規模は国により異なるが、栽培工程は同じである。エダマメの栽培工程は、大きく分けて5段階ある（図1-3）。

- ① 圃場への土壌改良資材・基肥の散布，耕うん
- ② 播種
- ③ 中耕・培土，殺虫剤や雑菌剤などの薬剤散布
- ④ 収穫
- ⑤ 出荷調製（脱さや，洗浄，外観精選別，袋詰め），出荷

エダマメはダイズ栽培用の機械を流用できることから，出荷調製以前の工程の機械化が進んでいる。しかし，ダイズとエダマメは収穫物が異なっており，ダイズは種子のみを，エダマメは茎全体もしくはさやのみを収穫する。そのため，ダイズ用の機械は収穫以降の作業ではエダマメへ使用できず，収穫以降の作業にエダマメ専用の機械を使用する。



図1-3 エダマメ栽培の各工程の作業光景

(2) エダマメの出荷調製について

エダマメ栽培でボトルネックになる工程は、出荷調製である。各作業の作業能率を表 1-1 に示す¹¹⁾。出荷調製の作業能率はそれ以前の作業よりも低く、いずれも 1.0a/h を下回っている。

エダマメの風味は、脱さや後に 20℃以上の常温で保管すると、全糖や遊離アミノ酸、アスコルピン酸の含有量低下により劣化する¹²⁾。そのため、出荷調製は風味の劣化防止の観点から迅速に行う必要がある。出荷調製は脱さや、洗浄、外観精選別、袋詰めを行う作業であり、収穫後に短時間で行う必要がある。栽培規模が大きい生産者は、人手を確保して出荷調製を短時間で行っている。

表 1-1 エダマメ栽培での各作業の作業能率

作業内容	作業機械	作業能率 (a/h)
耕うん・整地	バーチカルハロー	58.8
播種	直播式播種機	37.8
中耕・培土	ロータリー式中耕機	26.2
	ディスク式中耕機	71.2
薬剤散布	ブームスプレーヤー	141.4
収穫	ハーベスタ	3.2
脱さや	脱さや機	0.67
粗選別	振動式選別機	0.71
手選別（精選別）	人力	0.21

※引用:日本農作業学会 (https://jsfwr.org/fw_data/index.html)

(3) 外観精選別作業の課題について

外観精選別（以下、精選別）は、出荷調製で最も作業能率の低い作業である。精選別は、さやの形状や色、病虫害の有無などを確認してエダマメを良品と不良品へ分別する作業である。一粒や未熟さやなどの小型の不

良品の選別は、振動でふるい落とす粗選別機が用いられている。しかし、粗選別機はさや形状の不良や病虫害などの不良品を選別できない。そのため、精選別作業では、粗選別機で選別したエダマメを手選別することが主流である。手選別の作業能率は 0.21a/h と低いため、手選別は一般に複数人で行われる。このような背景から、手選別は、生産費の中の人件費の割合を増やす要因となっている。

ここで、秋田県仙北地域の例をあげる。この地域のエダマメ農家は 1 日に約 3.0 人を雇用して出荷調製を行っており、労働費が生産費のうち 46% も占めている¹³⁾。一方、生産費のうち、機械関係費の占める割合は 11% であり、労働費よりも低い割合である。しかし、この例は、精選別機を導入していない事例である。もし生産者が高価な精選別機を導入した場合、労働費は下がるが、機械費が上がると想定される。この場合、筆者は、機械費が増加するが、精選別機を導入すべきと考えている。その理由は、今後予想される労働者不足である。日本の農業就業人口は、2015 年から 2022 年の間に約 175 万人から 122 万人に減少した¹⁴⁾。そのため、継続的な雇用労働の確保は、現状の雇用を確保できているとしても、今後難しくなると予想される。これに対して、生産者は、精選別機を導入することで、将来予測される労働者不足の影響を少なくでき、安定的な生産を行えると考えられる。また、生産者は、労働費を削減しつつ精選別の作業能率を改善できる。今後生産者が精選別機を導入するには、開発される精選別機が手選別と同等以上の選別精度と低価格になる必要がある。

第3節 エダマメの精選別に関する研究

精選別機に必要な機能の1つに、エダマメの精選別アルゴリズムがある。精選別アルゴリズムの先行研究は、画像処理を利用した事例が報告されている。

佐々木ら¹⁵⁾は、デジタルカメラで撮影したエダマメ画像を二値化して周囲長などの形状特徴量(規則性)を算出し、それらの特徴量を遺伝的プログラミング(General Programming)と組み合わせて良品と欠粒さやを識別した。この選別方法は、良品と欠粒さやの2項目のみ判別可能であった。実際の精選項目は欠粒さや以外にも粒数やさやの変形、変色など複数あるが、それらには対応していなかった。

徳田ら¹⁶⁾は画像処理を用いた形状解析アルゴリズムを開発し、さやの湾曲や未熟粒といったさやの形状を選別した。この研究では、デジタルカメラでエダマメを上面と側面から撮影し、二値化した画像からさやの長軸長や短軸長、中心線長などの特徴量を計測して閾値分類した。このアルゴリズムは、未熟さや・欠粒さやを98.8%の精度で、湾曲さやを100%の精度で分類した。しかし、このアルゴリズムは、さやの形状に関する特徴量を計測して、さやの形状を解析するものであった。そのため、この研究は、さやの色や虫害などの色に関する選別を行っていない。そして、この研究では、エダマメを上面と側面の2方向から撮影して特徴量を取得した。この手法は選別レーンにエダマメが重ならないように1つずつ流す制約が発生し、処理量を増大させるには選別レーンを増やす必要がある。

片平ら¹⁷⁾は色彩情報を利用した画像処理とファイバーセンサを利用したエダマメの精選別機を開発した。試作機は、カメラ映像と画像処理により色やさやの形状を、ファイバーセンサでさやの厚みを検出することが可能であった。各検出部の直前には搬送ガイドが設置されており、これはエダマメを1個ずつ検出部に通す役目がある。試作機の作業能率は手選

別と比較して 1.8 倍から 5.1 倍に向上し、精選別の機械化の有効性を示した。しかし、子実熟度の検出に関し、エダマメが搬送ガイド通過時に不安定な姿勢となり、正確に識別できないことがあった。また、傷害検出に関し、画像処理ソフトウェアが、照明の陰影により誤判定することがあった。試作した選別機は、これらの原因で精選別精度に課題が残った。

以上より、先行研究の精選別アルゴリズムは、単体のアルゴリズムでさやの形や色、未熟度の全項目を選別できない課題があった。また、機構を追加して搬送部を流れるエダマメの向きを正確に揃える必要があった。高精度な精選別機を開発するには、これらの課題を解決できる従来手法と異なる新たなエダマメの外観精選別技術の確立が求められる。

第4節 ディープラーニングと AI

(1) 第3次 AI ブームと CNN の基本構造

近年、人工知能(Artificial Intelligence, 以下, AI)が、目覚ましい進化を遂げている。現在は第3次 AI ブームと呼ばれ、背景にはビッグデータと呼ばれる大量のデジタルデータの集積や CNN (Convolutional Neural Network : 畳み込みニューラルネットワーク) 構造の AI の発展, コンピューターの計算能力の向上, そしてビッグデータから特徴量を自動で学習するディープラーニング(Deep Learning, 以下, DL)¹⁸⁾の開発がある¹⁹⁾。

CNN の基本構造を図 1-4 に示す。CNN は、入力層と畳み込み層、プーリング層、全結合層、そして出力層を重ね合わせた構造となっている。CNN は、入力データから対象物の分類に必要な特徴量を抽出して分類結果を出力する²⁰⁾。ここで、各層の特徴を解説する。入力層は、入力データを CNN のサイズに合わせてリサイズして、ネットワークへ入力する。畳み込み層は、画像データに複数のフィルタ(カーネル)を用いて積和演算し、特徴量を抽出して特徴マップを作成する(図 1-5)。プーリング層は、特徴マップを特定サイズに分割して決められた演算を行う(図 1-6)。プーリング処理は、主に max プーリングもしくは avg プーリングが用いられる。max プーリングは、分割した各特徴マップの中からの最大値を抽出してダウンサンプリング画像を作成する。avg プーリングは、平均値を抽出してダウンサンプリング画像を作成する。全結合層は、特徴マップを 1 次元に変換する。そして、Softmax 関数などの活性化関数(Activation function)を適用して最終出力値を確率分布ベクトル化して分類を行う。最後に、分類結果は、出力層から出力される。

CNN の原型は、1982 年に福島が考案したネオコグニトロンである²¹⁾。ネオコグニトロンは、人間の視野角の単純型細胞と複雑型細胞の 2 つの働きを模したものである。単純型細胞は画像の濃淡パターンを検出し、

CNN では畳み込み層として表現される。複雑型細胞は物体の位置が変動しても同一の物体とみなす機能があり，CNN ではプーリング層として表現される。LeCun らは，1998 年に畳み込み層とプーリング層をそれぞれ 2 層組み合わせた LeNet を開発し，これが現在の多層型 CNN の原型となった²²⁾。

DL は，CNN が正しい分類を行えるような特徴量を抽出できるまで，CNN のパラメータを自動で最適化する機械学習手法の 1 つである。CNN は多層化することでより複雑な特徴を表現できることから，これまで困難とされてきた分類タスクに利用することが期待されている。

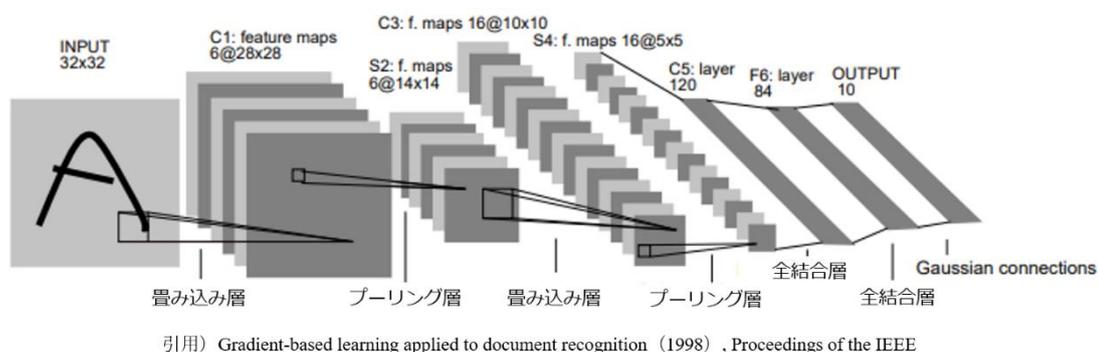


図 1-4 CNN (LeNet) の基本構造

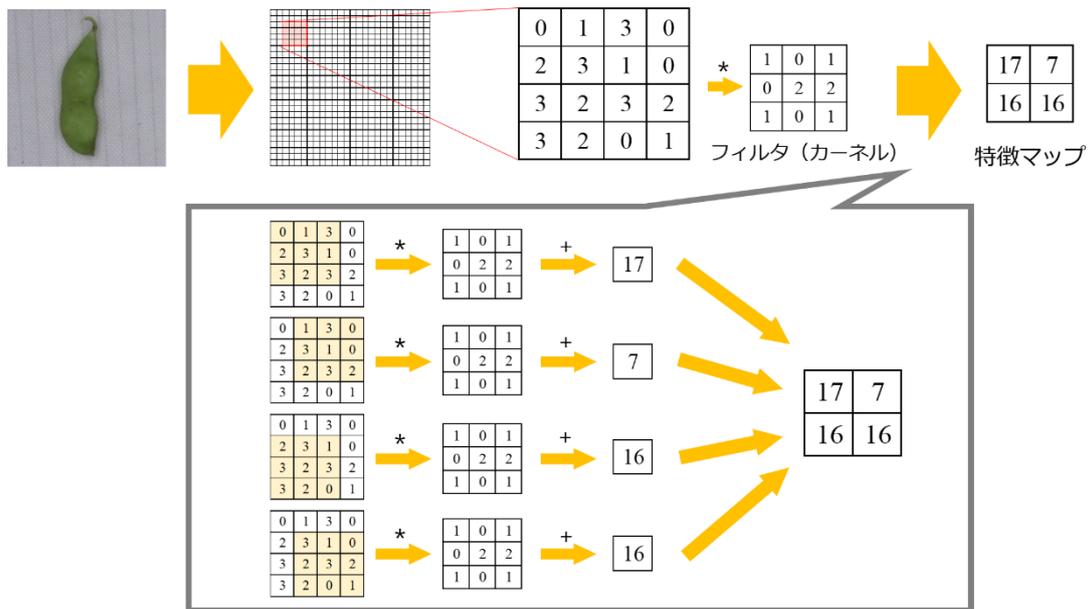


図 1-5 畳み込み層における積和演算による特徴抽出の例

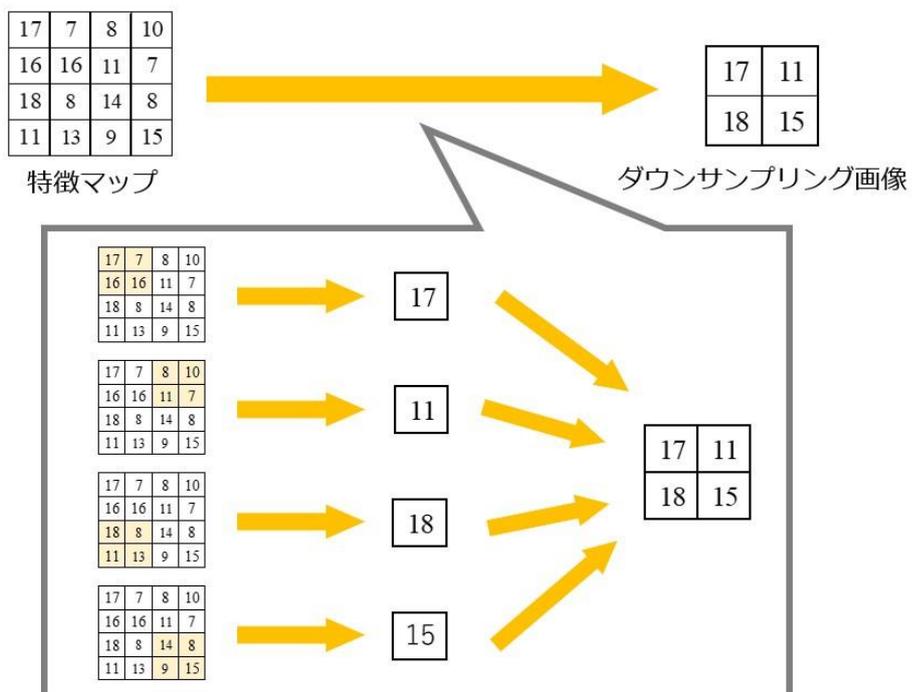


図 1-6 プーリング層における max プーリングの例

(2) 画像認識分野における AI の発展

AI にはいくつかの分野があり，その一つに画像認識がある。画像認識は，画像や動画に写っている対象物を分類するタスクである。画像認識に関し，ILSVRC（ImageNet Large Scale Visual Recognition Challenge）というコンテストが，2010 年から 2017 年にかけて行われた。ILSVRC は，ImageNet という学習データセットで AI を開発し，テストデータを 1000 クラスへ分類する精度を競うコンテストである。このコンテストで，AlexNet²³⁾が 2012 年に登場した。AlexNet は，DL で開発された最初の CNN 構造の AI であった。AlexNet の精度は 84.7%を記録し，過去の最高精度の AI の記録（74.2%）に大差をつけて優勝した。その後，様々な CNN が AlexNet を参考に開発され，2014 年には VGGNet²⁴⁾と GoogLeNet²⁵⁾が，2015 年には ResNet²⁶⁾が開発された。ResNet は 3.6%の画像分類エラー率を達成し，人間の 5.1%を超えたことが報告された²⁷⁾。すなわち，この時点で，AI は今まで人間が行っていた複雑な対象物の分類を行える可能性が示唆されたことになる。

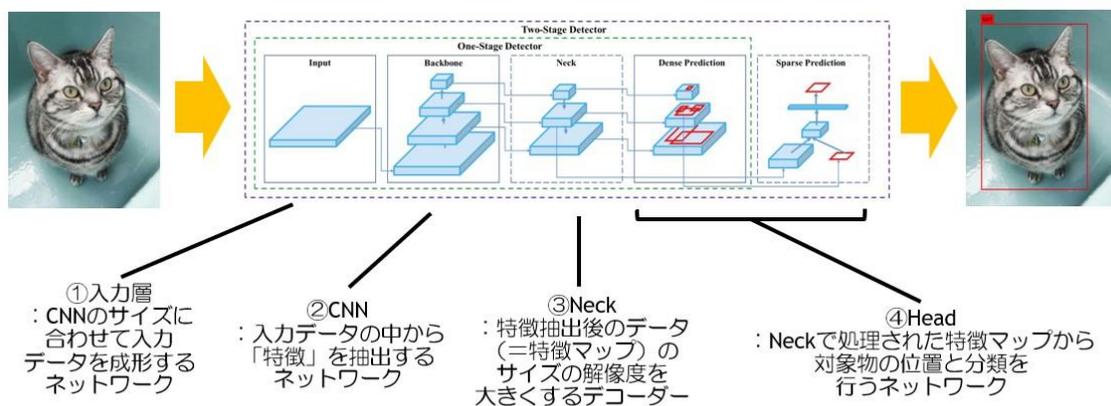
また，画像認識を発展させた物体検出も，CNN を応用することで進化している²⁸⁾。物体検出は画像や動画の中から「どこに，何が」あるのか探し出す技術であり，対象物の検出と分類の両方を 1 つのアルゴリズムで行える（図 1-7）。物体検出 AI は，AI に検出・分類させたい対象物が写っている画像や動画を使用して DL で開発される。代表的な物体検出 AI には，Faster R-CNN²⁹⁾や SSD³⁰⁾，YOLO³¹⁾，EfficientDet³²⁾などがある。

物体検出 AI は，主に 4 つのネットワークで構成される（図 1-8）。最初に，入力層が，入力データを CNN のサイズに合わせて変換する。次に，CNN が入力データから特徴量を抽出して特徴マップを作成する。そして，Neck 部が，その特徴マップをデコードする。最後に，Head 部が，対象物の領域推定と分類タスクを行う。物体検出 AI の精度や速度は，入力層の

サイズや CNN の種類， Head 部のアルゴリズムによって変化する。物体検出 AI は各ネットワーク構造の進化で検出・分類の精度と速度の向上が進んでおり，現在ではリアルタイムでの物体検出が可能となっている。



図 1-7 画像認識 AI と物体検出 AI の違い



引用：YOLOv4: Optimal Speed and Accuracy of Object Detection(2020), <https://arxiv.org/pdf/2004.10934.pdf>

図 1-8 物体検出 AI の基本構造 (YOLOv4³³⁾)

(3) DL で開発した AI の応用事例

DL で開発された AI の性能は日々向上しており、各分野への応用が期待されている。画像認識 AI や物体検出 AI は、農業分野でも応用研究が始まっている。上垣ら³⁴⁾はトマトの自動収穫機の開発、朝田ら³⁵⁾はロボットコンバインでの人との衝突回避、Apolo ら³⁶⁾はオレンジの収量予測といった研究に応用している。

また、画像認識 AI や物体検出 AI が、農作物の外観選別へ応用された事例も報告されている。Unay³⁷⁾はハイパースペクトルカメラで撮影したリンゴの等級選別を画像認識 AI で実施し、良品と不良品の 2 項目へ 95.6% の精度で分類した。Chen ら³⁸⁾は柑橘を良品と機械的損傷、病変の 3 項目へ分類する物体検出 AI を開発し、高い検出・分類精度を達成した。Bhupendra³⁹⁾らは EfficientNet⁴⁰⁾などの複数種類の画像認識 AI を開発し、インディカ米を 7 項目へ高精度に選別した。これらの事例より、今まで人が目視で行ってきた多くの作業は、AI で代替できるといえる。

物体検出 AI は、一度に複数の対象物を検出して複数の項目へ分類できる。さらに、対象物が映像内に様々な向きや位置にあっても検出・分類できる。よって、物体検出 AI をエダマメの精選別へ応用した場合、搬送部上に流れたエダマメの中から不良品を検出して、座標情報を特定することができる。そして、選別機のエジェクター（不良品除去部）へ動作信号を送り、不良品を除去できる。また、物体検出 AI は複数の選別項目へ分類可能であることから、物体検出 AI は生産者の細かな選別ニーズへ対応できる。さらに、物体検出 AI はカメラの映像だけで選別処理をできるため、搬送部や検出部の構造を簡素化して製造コストを抑制できる。

以上から、DL で開発した物体検出 AI は、エダマメ精選別機の精選別用ソフトウェアとして有用といえる。

第5節 DLで開発した物体検出AIの課題

(1) DLと従来の画像処理の違いについて

DLは、従来の画像処理と比べて、開発の方法や難易度などが異なる。DLと従来の画像処理の違いを表1-2に示す。

表 1-2 DLと従来の画像処理の違い

	DL	従来の画像処理
特徴抽出器の開発	自動	人力
検出・分類器の開発	自動	自動・半自動
パラメータ数	多	多
開発時間	条件により変動	条件により変動
開発難易度	易	難
パラメータの手動変更	難	易
AIの解釈の説明性	難	易



図 1-9 従来の画像処理（上段）と DL（下段）の各ネットワーク設計の違い

DLの最大の利点は、特徴抽出器のパラメータを自動調整できることである（図 1-9）⁴¹⁾。画像から対象物を見つけて分類するためには、まず特徴抽出器が、入力データから特徴量を抽出する。次に、検出・分類器がそ

の特徴量を使って検出・分類する。従来の画像処理に関し、検出・分類器の設計は、SVM (Support Vector Machine, サポートベクトルマシーン)⁴²⁾などの機械学習を利用して自動で行える。しかし、特徴抽出器の設計は、人力で行っている。人力での特徴抽出器の設計は時間がかかり、さらにプログラミングスキルが必要である。また、条件によっては、特徴抽出器の能力は、DLの方が高いとされる。

人力で開発した特徴抽出器に関する先行研究に、農地の果実や花の個数カウントがある。Aggelopoulou ら⁴³⁾は、果樹園で撮影した RGB 画像からリンゴの花の量を推定する画像処理技術を開発した。Aggelopoulou らの開発手法は、最初に RGB 画像をグレースケールへ変換し、さらにグレースケール画像を二値化して白い花を抽出した。Aggelopoulou らは、リンゴの白い花の特徴抽出を行いやすくする工夫として、樹体の背景に黒色の垂れ幕をかけた。これにより、アルゴリズムが、花の白色の画素を識別しやすくなった。しかし、この開発手法は、黒色の垂れ幕を設置した条件下のみで機能する問題点があった。Lin ら⁴⁴⁾は、RGB 画像に深度情報を加えた RGB-D 画像をもとに、柑橘の果実を検出する画像処理技術を開発した。Lin らの開発手法は、まず深度センサで取得した樹体までの距離情報をもとに、RGB 画像から果実や葉の領域を抽出し、得られた特徴量から SVM で果実の領域を特定する。この手法により、柑橘の検出は高精度に行われたが、果実同士が隣接している場合は複数の柑橘が 1 個体として誤認識されてしまうこと、深度センサの深度精度が太陽光の影響で低下し、果実を抽出できないことがあった。以上から、人力で開発した特徴抽出器はその労力に対して汎用性が低く、特定条件下でその性能を発揮するという課題があった。

一方、DL は、従来の画像処理とは異なり、特徴抽出器と検出・分類器を全自動で開発できる。また、CNN の特徴抽出方法は、従来の画像処理

のようなグレースケールへの変換や二値化を行うのではなく、入力データへの複数のフィルタを用いた畳み込み処理などを経て行う。DLは、学習データに様々な条件下で収集したデータを混合することで、共通する特徴量を抽出できるように各パラメータを自動調整する。これにより、CNNは、様々な条件下でも対応できるようになる。DLで開発したAIの汎用性の高さの事例として、Apoloらが開発した柑橘の検出AIがある。Apoloらは、マルチコプターで撮影したRGB画像から柑橘を検出・分類する物体検出AIを開発した。この研究では、物体検出AIが、様々な画像に写っている果実のほとんどを検出できたと報告している。

そして、DLは、開発の容易さという特徴もある。DLによる画像認識AIや物体検出AIの開発フローを図1-10に示す。最初に、開発者は、AIに検出・分類させたい対象物や分類項目を考える。次に、対象物が写っている画像や動画を収集する。物体検出AIを作成する場合、データ収集後にアノテーション作業でデータ内の対象物の位置座標と分類ラベルを付与して学習データを作成する。最後に、学習データを用いてDLを行うことで、AIが完成する。現在、DLの開発用フレームワークがインターネット上で多数公開されており、そのほとんどが無料で利用できる。DLの実行やAIの稼働はパソコン上で短いコマンドを実行するだけで容易に行えるため、プログラムの知識がない初心者も容易にAI開発を行える。

以上から、DLはAIの高精度化とAI開発の容易化を促進し、従来手法で精度の低かったタスクの発展やAIの新規利用の促進が期待される。

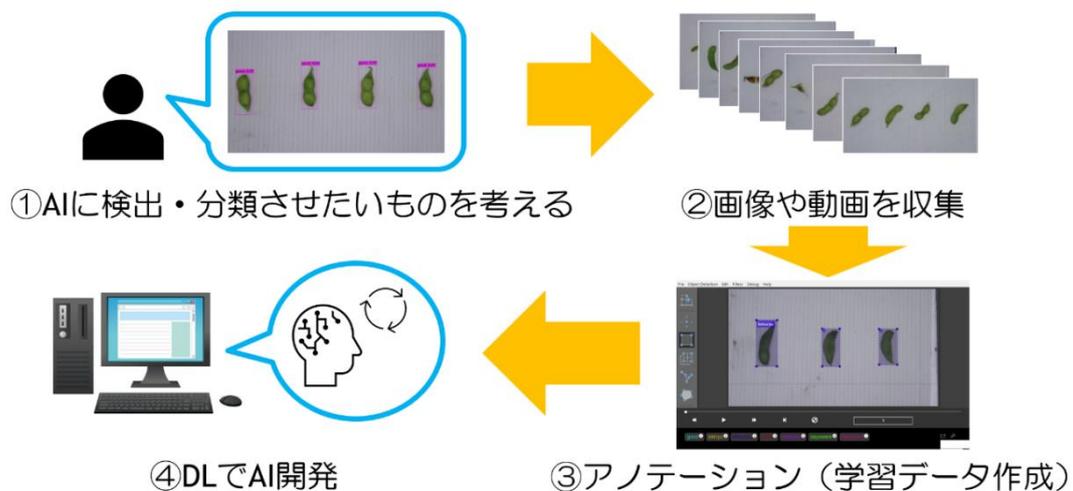


図 1-10 DL による画像認識 AI や物体検出 AI の開発フロー

(2) DL における学習データの研究について

DL で開発した AI の精度は、主に CNN の特徴抽出能力や DL の方法、そして学習データの内容に依存する。AI の精度は年々向上しており、2017 年に ILSVRC が終了した後も、ResNeXt⁴⁵⁾や EfficientNet といった改良モデルが多数報告された。

また近年、転移学習⁴⁶⁾が、新しい DL の方法として活用されている。転移学習は、事前に調整された AI モデルをベースとし、新たな学習データで AI を再学習させる手法である。転移学習で再学習される AI は、ImageNet や MS COCO⁴⁷⁾データセットのような数百万もの大規模なデータセットで開発されたものが利用されている。大規模データセットで開発された AI は、大量のデータで調整されており、特徴抽出パラメータの汎用性が高い。そのため、転移学習は、用意した学習データが少ない場合でも、高精度な AI の開発を可能とする。転移学習の有無による AI の精度比較例として、鄭らは、鶏の肝臓の健康度を分類する VGG16 の画像認識 AI を開発した⁴⁸⁾。この研究で、画像認識 AI は、転移学習により高精

度化することが報告された。転移学習以外の DL の新手法に、「データ拡張 (Data augmentation)」⁴⁹⁾がある。データ拡張は、元画像の回転や色情報の変換を行い、学習データ数を増やす手法である。データ拡張の有効性に関し、Suchet ら⁵⁰⁾は、果樹園内の果実を検出するための AI 開発にデータ拡張を導入した。この研究で、果実画像を回転するデータ拡張が、果実の検出精度の向上に寄与したことを報告している。このように、多くの研究者が、AI の構造や DL の手法について活発に研究を行っている。しかし、AI の根幹となる学習データの内容に関する研究は少ない。

この理由は、AI の開発者が 2 つのタイプに分かれているためと考える。1 つ目は、AI の構造や DL の手法を考える研究者 (以下、タイプ 1) である。2 つ目は、現場のニーズに応えた AI を開発する研究者 (以下、タイプ 2) である。現在、AI の開発者は、タイプ 1 が大半である。その理由は、ImageNet や MS COCO などのオープンデータセットが整備されており、同じデータを用いて開発した AI の精度比較が容易であるためである。一方、タイプ 2 は、各自が異なるデータを収集し、用途に合った AI 開発を行っている。そのため、タイプ 2 の研究者は、データや用途の違いにより、他の研究との比較を行い難い。また、タイプ 2 による先行研究は、従来の画像処理と DL で開発した AI の精度比較や、DL の有用性の確認に焦点が当てられていることが多い。これらの理由により、タイプ 2 の研究事例は特定の事例に多いが、エダマメの精選別に関する事例はほとんどない。

高精度な AI 開発に必要な学習データの内容が分かっていない場合、データの収集方針が定まらない。また、AI の精度低下が確認された場合、その原因が学習データとアルゴリズムのどちらであるか特定が難しい。近年、MLOps (Machine Learning Operations)⁵¹⁾が、AI の開発フェーズと運用フェーズにおいて重要となっている (図 1-11)。MLOps は、AI の品

質保持と継続利用を目的とした仕組みである。MLOps では、開発した AI を運用中も継続的にモニタリングし、システム運用や使用環境に応じて AI を再学習する。これにより、AI は、安定的かつ長期的に運用されるようになる。AI の再学習が必要となった場合、開発者は、新たな学習データを用意する。しかし、開発者は AI の挙動を確認して、どのような学習データが過不足していたか判断する必要がある。AI 開発で最も時間と労力がかかる作業はデータ収集と整備とされ、AI 開発のボトルネックとなっている⁵²⁾。事前に学習データの内容とそれに対する AI の精度変化を明確化できれば、開発者はデータ収集や整備をより効率的かつ効果的に行える。よって、学習データの内容研究は、AI 開発の円滑化や AI の精度改善・精度保証につながり、最終的に AI の利用促進や各種課題解決につながる。エダマメの精選別では、物体検出 AI を搭載したエダマメ選別機が開発され、精選別の作業能率改善が見込まれる。そのため、DL でエダマメ精選別用の物体検出 AI を開発するには、学習データ内のどのような要素が精度変化に影響するか調査し、それに応じた学習データの構築方法を考える必要がある。

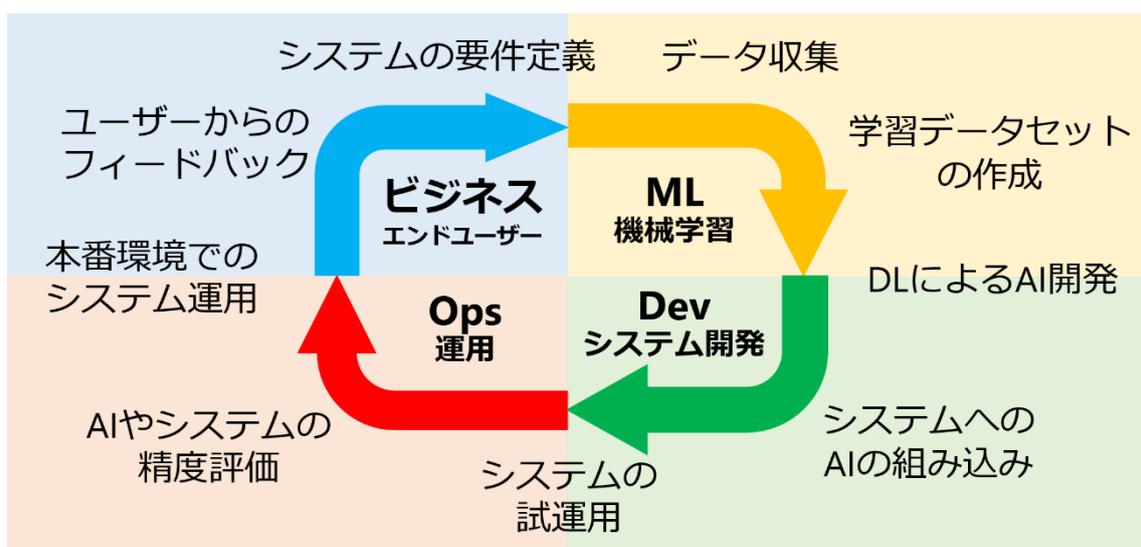


図 1-11 MLOps による AI の開発フェーズから運用フェーズまでのフロー

第 6 節 研究の目的と目標

本研究の目的は、エダマメ精選別作業の省力化と生産者の収益を向上させることである。そして、目標は、DL で高精度なエダマメ精選別用の物体検出 AI を開発することである。さらに、エダマメ精選別用の MLOps を考案する。本研究では、最初に DL でエダマメ精選別用の物体検出 AI の開発と精度評価を行った。この研究で、物体検出 AI の学習データセットの構成を変化させ、学習データの内容の変化による精度変化の整理と、その原因を調査した。最終的に、高精度な精選別用の物体検出 AI 開発に必要な学習データの構築要素や物体検出 AI の開発体系を検討した。

第 7 節 論文の構成

本論文は第 1 章の序言から始まり、DL でエダマメ精選別用の物体検出 AI の開発から、学習データの構築方法、AI の開発体系について検討した全 4 章で構成されている。第 2 章では、品種の組み合わせや CNN、物体検出 AI の種類を変えた物体検出 AI を開発し、手選別との外観精選別精度の比較や、学習データに含むエダマメ品種の違いが物体検出 AI の精度に及ぼす影響を調査する。第 3 章では、CNN の分類根拠可視化技術「Score-CAM」を使用し、学習データに含むエダマメ品種の違いが AI の精度変化に及ぼす原因を調査する。第 4 章では、総合考察として、学習データの観点から AI 開発体系の提案や、精選別用の AI の選別以外への応用を検討し、本稿を終了する。

第Ⅱ章

学習データセットに含まれるエダマメ品種の違いが 物体検出 AI の精選別精度に与える影響

第 1 節 緒言

(1) 学習データの内容が AI の精度に影響した例について

DL は学習データが多いほど、対象物の特徴を学習して高精度な AI を開発できるため、複数の開発者が同じ用途の AI を開発する場合に学習データを大量に集められるメリットがある。一方、デメリットは、異なる条件で収集された学習データによる AI の精度低下である。これは、学習データが収集者の個々のルールに基づいて収集されることに起因する。例として、PlantVillage データセット⁵³⁾の事例がある。PlantVillage は、2015 年に公開された植物病害の画像分類用オープンデータセットである。そして、多くの一般人が、PlantVillage のデータ収集に参加した。そのため、PlantVillage には、屋内外の様々な場所で収集されたデータが含まれている。Ferentinos らは、屋内で撮影された PlantVillage のデータで学習した画像認識 AI に屋外で撮影されたテストデータを分類させた結果、画像認識 AI の精度が 33%と低かったことを報告している⁵⁴⁾。この精度の低さは、学習データとテストデータの背景や光条件の違いによるものと考えられた。さらに、データセットの正解ラベルが本当に正しいか問題視され⁵⁵⁾、その後、PlantVillage は公開中止となった。このことから、学習データの収集と AI の運用は、統一ルールに基づいて行う必要がある。

このルールは、事前に学習データに起因する AI の精度低下要因を特定することで作成できる。AI の精度は精度評価を通じて初めて分かるため、開発者が目的とする AI を一度開発して学習データの問題点を帰納的に特定する必要がある。DL で開発した物体検出 AI をエダマメの精選別へ

応用した先行研究に関し、森ら⁵⁶⁾は 画像内からエダマメの検出と品質分類をする YOLOv3⁵⁷⁾の物体検出 AI を開発した。この物体検出 AI は、良品 1 項目と不良品 7 項目へ分類可能であり、人間と同等の良品・不良品分類精度を達成したことから、DL で開発した物体検出 AI がエダマメの精選別に有用であることを示した。そして、物体検出 AI の精度は、学習データの違いによって変化することが確認された。先行研究でのエダマメ画像は、屋内に設置したエダマメ精選別機で収集し、光条件や背景の色、そしてカメラとエダマメまでの距離はすべて統一されていた。一方、学習データセットは、品種の組み合わせを変えて 2 通り設定されていたため、この精度変化が画像に写っているエダマメの品種に起因すると考えられた。開発した物体検出 AI の精度は、庄内三号のテスト画像で精度評価した結果、2 品種を混合することでより低くなった。この結果から、精選別用の物体検出 AI は、学習データに選別対象の品種以外を含むと精度が低下する可能性がある。しかし、この先行研究では、庄内三号と秘伝の 2 種類のエダマメのみを使用し、品種の組み合わせを 2 通りのみ設定していた (図 2-1)。さらに、開発した物体検出 AI は過学習の可能性があったため、学習データに含まれる品種の違いが物体検出 AI の精度に及ぼす影響を明確にできなかった。過学習は、開発した AI が学習データに過剰適合し、新規の入力データに対する分類精度が低下する現象である。

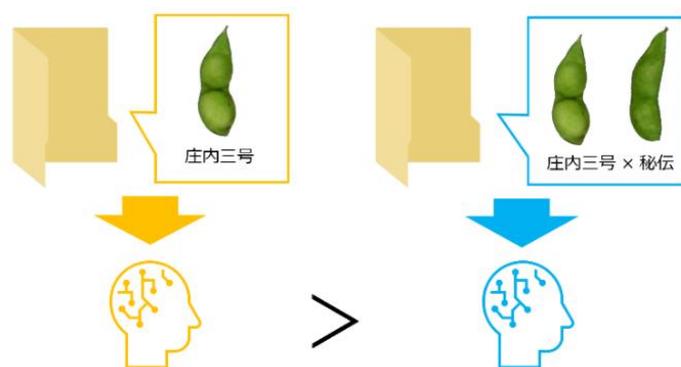


図 2-1 先行研究におけるエダマメ品種の組み合わせ方と AI の精度比較

農作物の外観選別へ DL を応用した先行研究は、いくつか報告されている。しかし、それらの先行研究では、複数の品種を利用した事例や意図的に品種の組み合わせを変えて学習データを構築した事例が少ない。そのため、品種が AI の精度変化に及ぼす影響は未解明なことが多い。

(2) ネットワーク構造の異なる物体検出 AI

また、物体検出 AI や CNN は、様々な種類がある。AI の利用者は 2 タイプ存在し、物体検出速度を重視するタイプと精度を優先するタイプがある。そのため、開発者は、各タイプに応じた構造の AI を開発する必要がある。物体検出 AI は、Head 部の構造によって Two-stage 型と One-stage 型の 2 種類に分けられる⁵⁸⁾。Two-stage 型は物体の候補領域の推定とクラス分類をそれぞれ 2 つの異なるネットワークで行い、主なモデルに Faster R-CNN や R-FCN⁵⁹⁾がある。One-stage 型はそれらを 1 つのネットワークで行い、主なモデルに SSD や YOLO, EfficientDet がある。One-stage 型の物体検出 AI は、Two-stage 型の物体検出 AI よりも物体検出速度に優れている。物体検出 AI の検出速度に関し、Koirala ら⁶⁰⁾は果樹園内の果実を検出する物体検出 AI を開発し、Faster R-CNN と SSD, YOLO の物体検出速度を比較した。その結果、512×512 ピクセルのテスト画像を使用した場合、物体検出速度は Faster R-CNN で 37~67 ミリ秒、SSD で 46~70 ミリ秒、YOLO で 15~25 ミリ秒であった。このことから、One-Stage 型の物体検出 AI は、物体検出速度が速いことが分かる。また、野村ら⁶¹⁾は、コンクリート表面のひび割れ検出システムに YOLO を適用した。この研究では、YOLO は 1,280×960 ピクセルの画像を 1 枚につき約 0.1 秒で処理し、リアルタイム性能の高さを示した。そして、森らは高速性が求められることを考慮し、エダマメ精選別用の物体検出 AI を YOLOv3 で開発した。このように、AI は要求性能に応じて今後様々な構造ものが開発され

ると予想される。

(3)本実験の目標

本実験は高精度なエダマメ精選別用の物体検出 AI を開発することを目標とした。本実験では、3 品種のエダマメを用いて品種の組み合わせを変化させた複数の学習データセットを設定した。そして、学習データに含まれるエダマメ品種の違いが、物体検出 AI の精度に与える影響を調査した。

第2節 実験方法

本実験のフローを図 2-2 に示す。最初に、エダマメを 7 通りに手選別して画像を収集し、それらを物体検出 AI（以下、本章では「AI」と称する）開発用と精度評価用に振り分けた。そして、AI 開発用データへアノテーションを行い、品種の組み合わせを変えた 7 つのデータセットを用意した。その後、DL で YOLOv3 と Faster R-CNN の 2 種類の AI を開発した。最後に、精度評価用のデータを各 AI へ入力して、精度評価を行った。

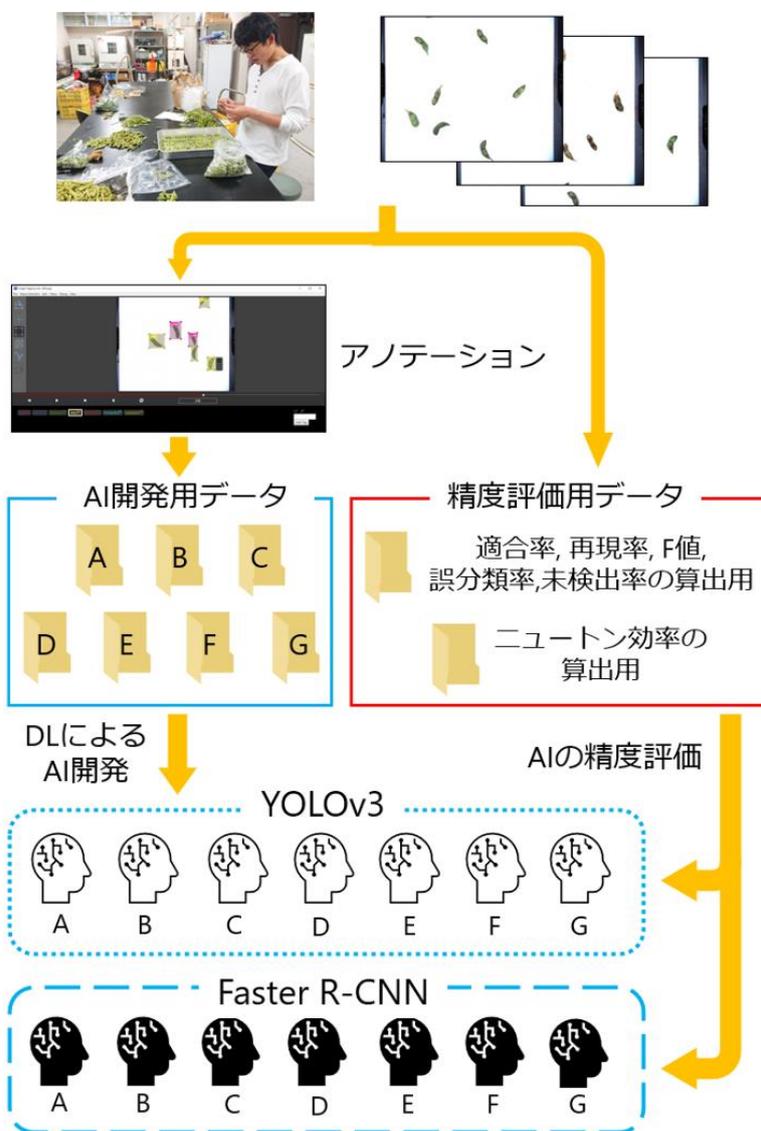


図 2-2 AI の開発から精度評価までの手順

(1) 供試材料と画像の収集

1) 本実験で使用したエダマメ品種と外観品質分類項目について

本実験で使用したエダマメは、2019年に山形大学農学部附属やまがたフィールド科学センター高坂農場で栽培した。品種は「庄内三号」と「庄内一号」,「グリーン75」の3品種である。庄内三号と庄内一号は、山形県庄内地方特産のダダチャマメ系統の茶豆である。グリーン75は、大さやの青豆品種である。本実験でこれらの品種を使用した理由は、物体検出AIの精度が外観の特徴が似ている品種と異なる品種を学習データに混合した場合にどのように変化するか確認するためである。エダマメは、収穫後に品種ごとに良品1項目と不良品6項目の計7通りに手選別した(図2-3)。

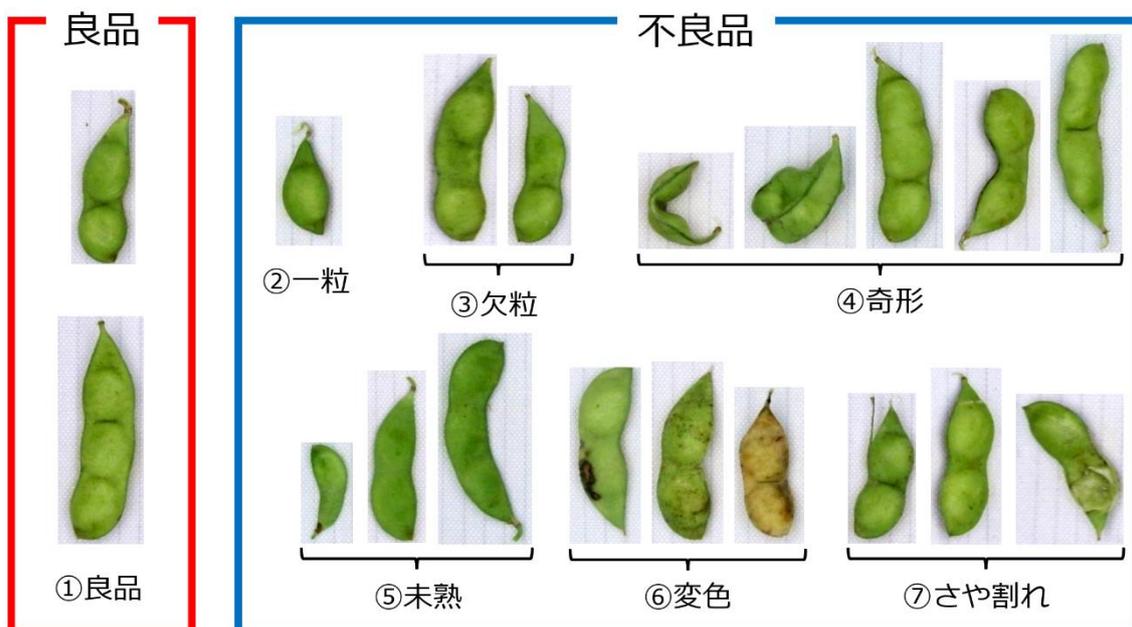


図 2-3 7項目のエダマメ分類項目

7項目の外観品質分類項目(以下、分類項目)は、次の基準で分類された。良品はさやに子実が2~3粒入っており、子実部のさやの厚みが8mm以上の個体である。また、良品は、さやにしみや虫害、変形が発生してい

ない個体とした。一粒はさやに子実が 1 粒のみ入っており、子実部のさやの厚みが 8 mm 以上の個体とした。なお、一粒で厚みが 8 mm 未満の個体は、未熟に分類した。欠粒は 2 粒以上の子実が入るさや形状で、子実部のさやの厚みが 8 mm 以上と未満の子実が混在している個体とした。奇形は、良品に比べて、さやの形状が異なる個体や子実が 4 粒以上含まれる個体とした。未熟は、子実部のさやの厚みが全て 8 mm 未満の個体とした。さや割れは、さやの一部にすり傷や割れ、破裂、欠けがある個体とした。変色は、さやの一部に黒色の斑点や模様がある個体、虫の食害がある個体、登熟が進んで黄色く変色している個体とした。なお、エダマメは、2 つ以上の分類要素が含まれる場合、選別者の主観によって特徴の大きい方へ分類した。

2) 本実験で収集した画像について

本実験の画像は、市販のエダマメ精選別機（株式会社ガオチャオエンジニアリング、PITA-EDS-mini01）を使用して撮影された。筆者は、エダマメを分類項目ごとにランダムに精選別機へ投入し、タイムラプス機能を用いて画像を収集した。タイムラプス機能を用いた理由は、短時間で画像を収集して、エダマメの変色を防ぐためである。画像のサイズは 1,280×1,024 ピクセルであり、データ形式は JPEG である。各画像には、複数のエダマメが不規則な角度で写っている（図 2-4）。本実験では、大量の画像を収集するため、精選別機の排出口から出てきたエダマメを投入口に再度投入して、再撮影した。この方法は同じエダマメを何度も撮影することになるが、エダマメの角度や位置は毎回変化しながらベルトコンベア上を流れる。これにより、エダマメはすべて異なる写り方になるため、学習データとして問題ないと判断した。撮影した画像は AI 開発と精度評価の両方に使用した。

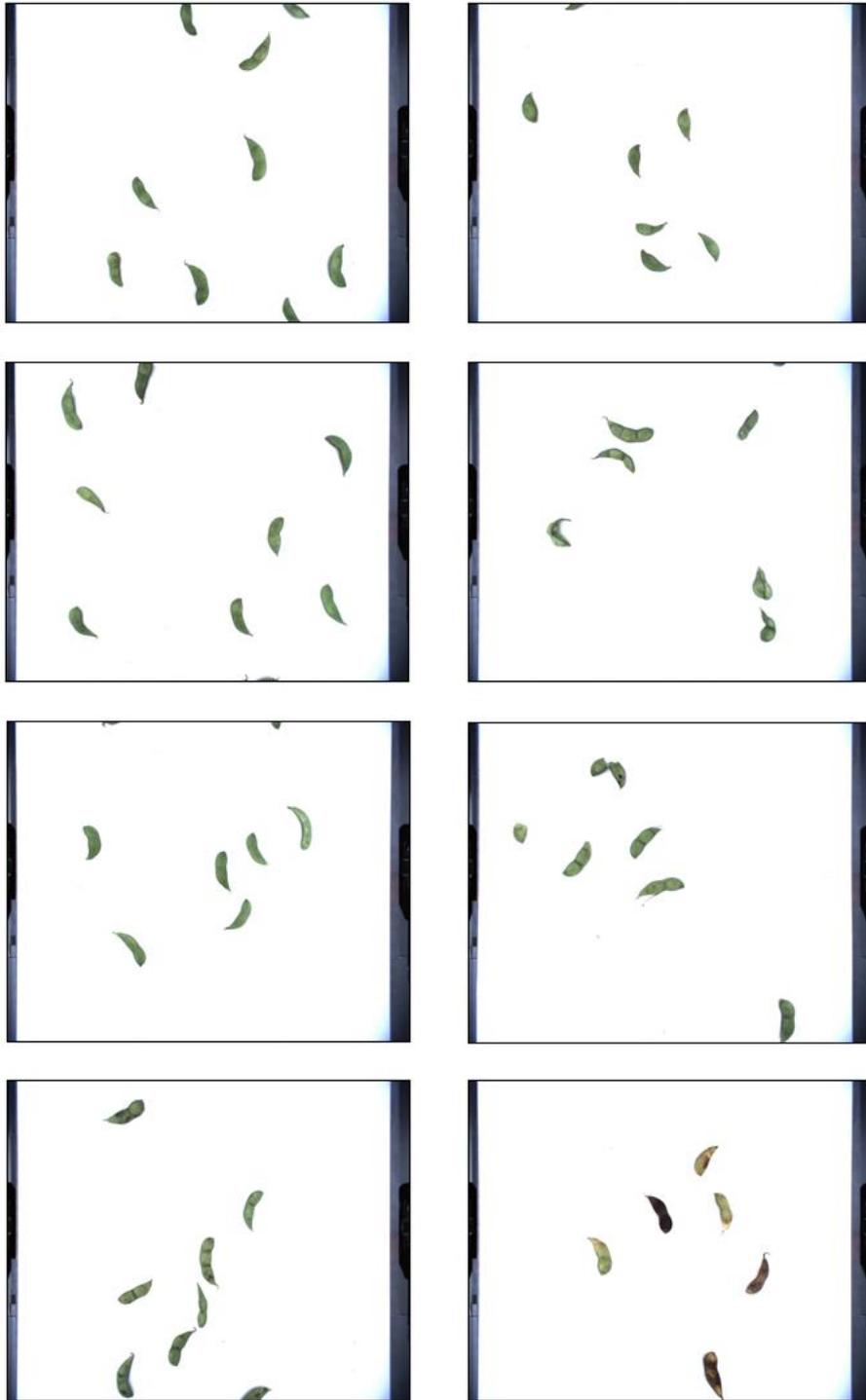


図 2-4 本実験で収集したエダマメ画像例

(2) アノテーションによる教師データの作成

AIに検出・分類させるエダマメは、図 2-5 に示す 7 項目である。森らの先行研究から、精選別用の AI の精度は学習データに良品と不良品の両方を含めることで、より向上することが確認された。この結果より、本実験のすべての学習データには、良品と不良品の両データを含めた。

筆者は画像収集後に AI 開発用データのアノテーションを行った。アノテーションは、画像内のエダマメの範囲指定と分類項目のラベルをエダマメ 1 つ 1 つに付与する作業であり、Microsoft 社のフリーソフトウェア「VoTT-1.7.2」で行った（図 2-6）。アノテーション範囲は、図 2-5 に示すオレンジ枠のとおりである。アノテーション範囲は 2 パターンあり、1 つ目がさや全体、2 つ目がさやの一部を範囲指定した。さや全体をアノテーションした分類項目は、良品と一粒、奇形、未熟、変色の黄化、さや全体にしみやさや割れが発生している個体である。一方、さやの一部をアノテーションした分類項目は、欠粒や、部分的なさや割れ、しみと虫害がさやの一部に発生している個体である。これらの分類項目は傷害部分を除き良品とほとんど同じ外観をしており、さや全体をアノテーションすると AI の精度を低下させる可能性があると考え、アノテーション範囲を限定した。

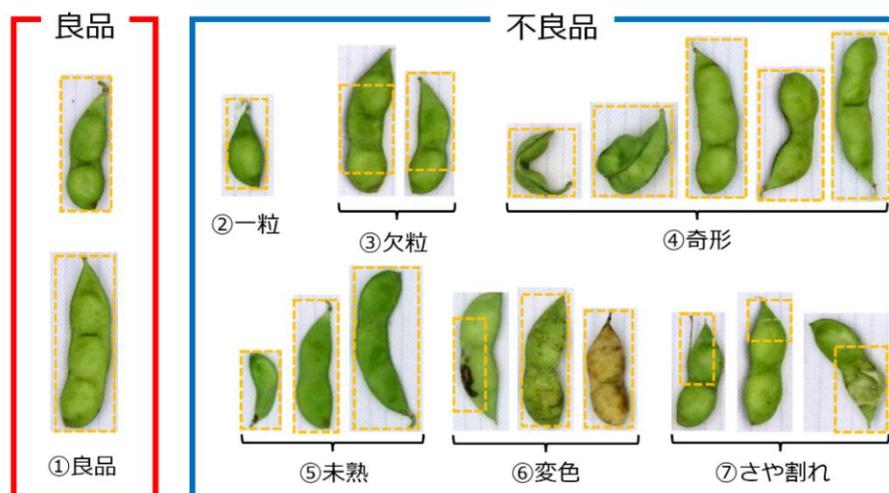


図 2-5 AIに検出・分類させる分類項目とアノテーションでの指定範囲

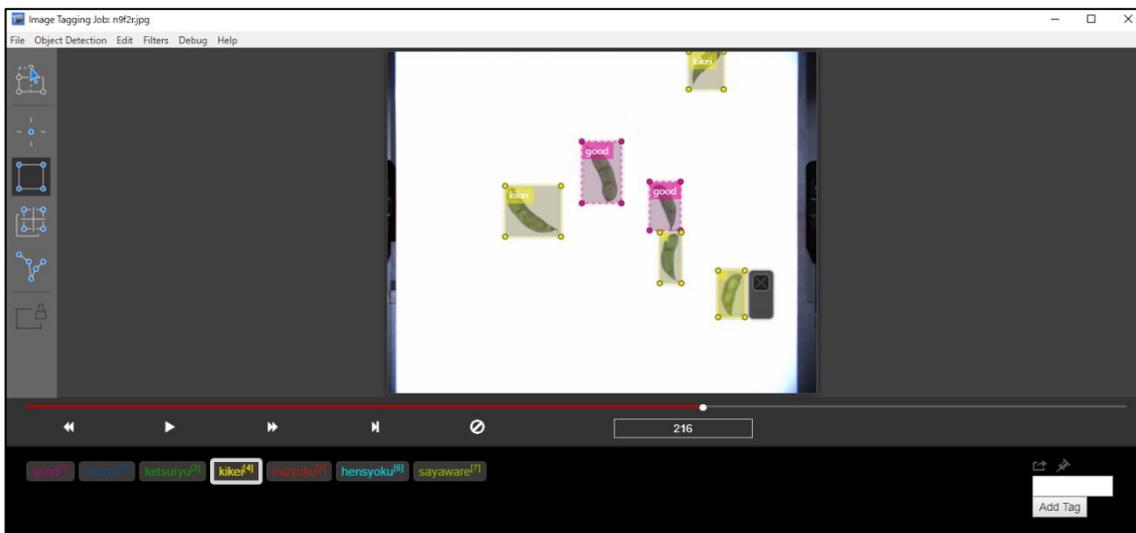


図 2-6 VoTT-1.7.2 によるエダマメ画像への
アノテーション例

学習データのアノテーション数と画像数を表 2-1 に示す。各分類項目のアノテーション個数は完璧に揃えづらく、そこで、アノテーション個数が 1000 個以上になるように調整した。この原因は、画像内に写っているエダマメの写り方にあった。エダマメは分類項目ごとに撮影したため、画像内のエダマメは 1 種類の分類項目のみ写ると考えていた。しかし、いくつかのエダマメは、写り具合で別の分類項目へ分類すべきと考えられるものがあった。それらのエダマメは、本来予定していた分類項目のラベルを付けず、画像データの写り具合から判別した分類項目のラベルを付与した。

アノテーションに関し、一部のエダマメは上下に重なった状態で撮影された。それらのエダマメは、重なった状態でそれぞれをアノテーションした。また、画面の端に見切れているエダマメは、さやの半分以上が写っている場合にアノテーションをした。

表 2-1 教師データ用のエダマメ画像のアノテーション個数と画像データ数

外観品質 分類項目	アノテーション個数 (個)			画像データ (枚)		
	庄内 三号	庄内 一号	グリーン 75	庄内 三号	庄内 一号	グリーン 75
良品	1267	1698	1448	192	183	220
一粒	1138	1119	1089	190	167	188
欠粒	1180	1353	1160	240	242	322
奇形	1375	1357	1224	269	304	240
未熟	1235	1477	1145	151	136	220
変色	1113	1376	1153	166	163	264
さや割れ	1151	1371	1094	241	282	211
合計	8459	9751	8313	1449	1477	1665

(3) 品種の組み合わせを変えた7つのデータセットの構成

学習データに含む品種の違いがAIの精度にどのように影響するか調査するため、A～Gの7種類のデータセットを設定した(図2-7)。なお、庄内一号は黄化の画像がなかったため、庄内三号の黄化の画像で代替した。データセットAとB、Cは単一品種のみで構成し、データセットDとE、F、Gは複数の品種を組み合わせで構成した。

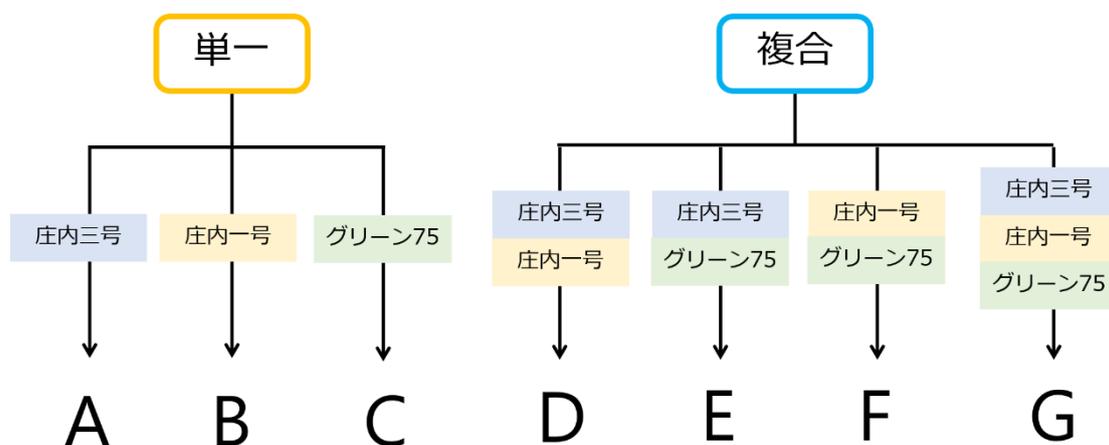


図 2-7 品種の組み合わせを変えた7種類のデータセットの構成

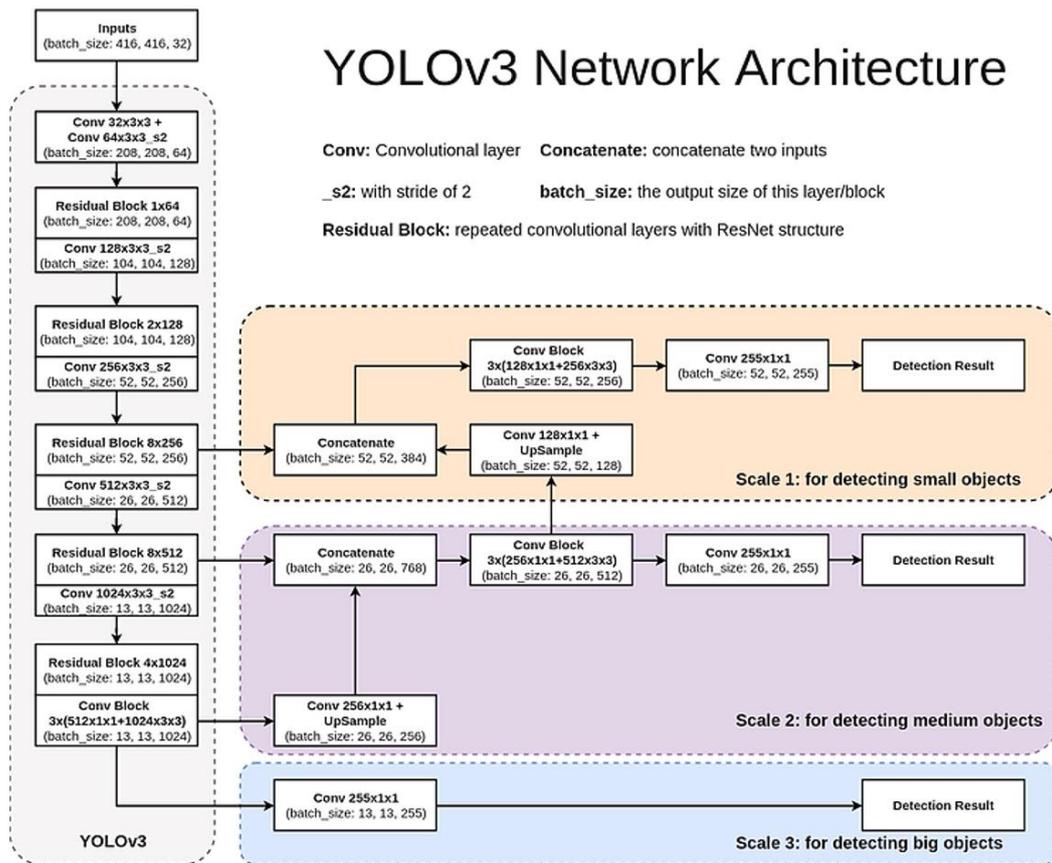
(4) DL による AI の開発

本実験で開発した AI は、YOLOv3 と Faster R-CNN の 2 種類である。これにより、学習データの品種の違いが、異なる構造の AI の精度に与える影響を調査した。

YOLOv3 と Faster R-CNN のアルゴリズムの模式図を図 2-8~2-9 に示す。YOLOv3 は、最初に入力層でデータを特定サイズに変換して、ネットワークへ入力する。次に、Darknet53 と呼ばれる 53 層の CNN が、特徴抽出を行う。そして、Head 部の FPN (Feature Pyramid Networks, 特徴ピラミッドネットワーク)⁶²⁾ がスケールの異なる 3 段の特徴マップの情報を抽出し、各特徴マップにアンカー (Anchor) を中心とした 3 つの異なるサイズのアンカーボックス (Anchor box) を配置する。これらのアンカーボックスは、特徴マップの中から対象物の領域推定を行う。最後に、Head 部は予測した推定領域内の特徴量を使用して、ロジスティック回帰により分類を行う。

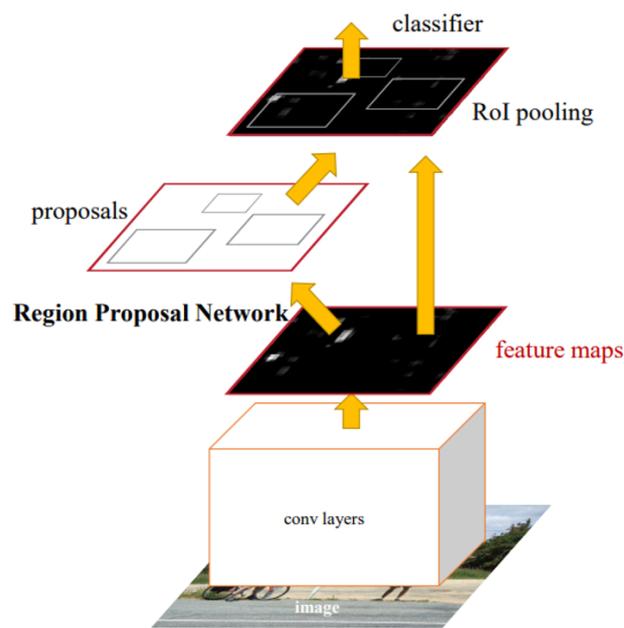
Faster R-CNN は、入力層と CNN までの構造が YOLOv3 と同様である。しかし、Faster R-CNN は、領域推定と分類の方法が異なる。Faster R-CNN は、CNN で抽出した特徴マップから RPN (Region Proposal Network, 領域提案ネットワーク) で対象物の領域推定を行う。RPN は特徴マップ上にアンカーを配置し、アンカーごとに 9 つの異なるサイズのアンカーボックスを生成する。対象物が各アンカーボックス内に含まれていると推定された場合、推定領域内の特徴マップが max プーリングで特定サイズへ変換される (RoI プーリング)。そして、推定領域の特徴量をもとに分類予測を行う。

YOLOv3 Network Architecture



引用 : DMP BLOG , 2018. What's new in YOLOv3, <https://blog.dmprof.com/post/a-closer-look-at-yolov3>.⁴⁶⁾

図 2-8 YOLOv3 の各ネットワークの基本構造



引用 : Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. P.3 Figure2.

図 2-9 Faster R-CNN の各ネットワークの基本構造

AI 開発に使用したフレームワークは、YOLOv3 で Darknet, Faster R-CNN で Tensorflow-gpu-1.13 (以下, Tensorflow) である。各 AI の CNN は、YOLOv3 で Darknet53 を, Faster R-CNN で ResNet101 を選定した。DL は、事前訓練済みモデルを利用して転移学習を行った。事前訓練済みモデルに関し、YOLOv3 は ImageNet データセットで、Faster R-CNN は MS COCO データセットで事前に学習されたモデルである。本実験で開発した AI の総数は、YOLOv3 と Faster R-CNN の 2 種類に対してデータセットが 7 種類であることから、合計 14 種類である。

DL 時のハイパーパラメーターは、学習回数を 30,000 回、学習データの Train と Validation の分割割合を 8:2、入力層のサイズを 768×768 ピクセルに設定した。学習率は初期値を 0.003 に設定した。さらに、学習率は学

習回数が 20,000 回と 25,000 回に達すると, 0.0003 と 0.00003 に減少するように設定した。Batch サイズは Darknet で 64, Tensorflow で 4 に設定した。これは, それぞれのフレームワークが DL で必要とする GPU メモリ量が異なるためである。AI の開発と精度評価に使用したコンピュータの環境を表 2-2 に示す。

表 2-2 DL と物体検出を行ったコンピュータの環境

OS	Ubuntu18.04 LTS
CPU	Intel Core-i7 7700K
GPU	NVIDIA GeForce GTX1080Ti
NVIDIAドライバー	410
CUDA	10.0
cuDNN	7.6.5
OpenCV	3.4.3
Python	3.6

(5) AI の精度評価

AI の精度評価は、6 つの指標で行った。精度評価の指標は、適合率と再現率、F 値、ニュートン効率、誤分類率、未検出率である。精度評価用の画像は、すべて庄内三号を使用した。

1) 適合率，再現率，F 値

AI が 7 項目のエダマメをそれぞれどれだけ正確に検出・分類できるか調査するため、本実験では適合率と再現率、F 値を算出して評価した。これらの指標は、次の式(1)~(3)で算出した。

$$\begin{array}{l} \text{適合率} \\ \text{(Precision)} \end{array} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\begin{array}{l} \text{再現率} \\ \text{(Recall)} \end{array} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{F値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}} \quad (3)$$

最初に、TP (True Positive : 真陽性) と FP (False Positive : 偽陽性), FN (False Negative : 偽陰性) について解説する。TP は、AI がテスト画像内のエダマメを正確に検出・分類できたケースである。FP は、AI が検出したエダマメの分類項目を誤分類したケースである。FN は、AI が検出対象のエダマメを未検出したケースである。精度評価者は、AI がテスト画像内の各エダマメをどのように検出・分類したか確認し、TP と FP, FN のいずれに該当するか確認した。そして、それらの個数を式(1)~(3)に代入

することで、分類項目ごとの適合率と再現率、F 値を算出した。

表 2-3 TP や FP, FN を判断するための混同行列

		AIの予測値	
		正	誤
実際の 正解値	正	TP (True Positive : 真陽性)	FN (False Negative : 偽陰性)
	誤	FP (False Positive : 偽陽性)	TN (True Negative : 真陰性)

適合率は、AI が検出した全てのエダマメのうち、正確に分類できた割合を示す指標である。適合率は、1 に近いほど、AI が誤分類していないことを示す。

再現率は、AI が検出・分類対象のエダマメを正確に検出・分類した割合を示す指標である。再現率は、1 に近いほど、AI が対象のエダマメを見逃していないことを示す。

F 値は、適合率と再現率の調和平均である。F 値は、1 に近い程、AI がエダマメを未検出・誤分類することが少ないことを示す。

適合率と再現率、F 値のテスト画像は、新たに良品と不良品 6 項目の画像を用意した。テスト画像は分類項目ごとに複数枚用意し、画像内のエダマメの合計数が分類項目ごとに 20 個になるように調整した。適合率と再現率、F 値の調査は 10 反復のテストデータを用意して行い、その平均値を求めた。

2) ニュートン効率

ニュートン効率は、AI がエダマメを良品と不良品の 2 項目へ正確に分類できた度合いを示す指標であり、農作物の選別精度指標の 1 つである

64)。本実験でニュートン効率を調査した理由は、AI と手選別の精度を比較するためである。

ニュートン効率の算出では、新たに良品と不良品の 2 項目のテスト画像を用意した。不良品のテスト画像は、不良品 6 項目の全てを均等に含む。テスト画像は複数枚用意し、画像内のエダマメの合計数が良品と不良品でそれぞれ 25 個になるように調整した。そして、ニュートン効率 (η) は、式(4)で算出した。

$$\eta \quad (\text{ニュートン効率}) = \frac{W_g}{25} + \frac{W_f}{25} - 1 \quad (4)$$

W_g は AI が良品を正しく検出・分類できた個数を表す。そして、 W_f は AI が不良品を正しく検出・分類できた個数を表す。ニュートン効率は -1 から 1 の値をとり、1 に近いほど、AI が正確にエダマメを良品と不良品へ選別できていることを示す。なお、ニュートン効率は重量ベースで評価する指標であるため、本実験ではすべてのエダマメの重量を同一と仮定して算出した。

AI による不良品の分類は、不良品のエダマメを不良品 6 項目のいずれかに分類できれば、正判定とした。ニュートン効率の調査は 10 反復のテストデータを用意して行い、その平均値を求めた。

3) 誤分類率と未検出率

誤分類率と未検出率は、AI がどのように誤分類や未検出をしたか確認する指標である。誤分類率と未検出率の調査に使用した AI は、適合率と再現率、F 値、そしてニュートン効率の結果を基に、最も精度が優れていると判断したものを使用した。この理由は、精度の高い AI でも分類や検

出が困難な分類項目を特定し、その結果を基に学習データの構築方法を検討するためである。

誤分類率と未検出率の調査は、適合率と再現率、F 値の調査をした時の誤分類数と未検出数を計測して算出した。適合率と再現率、F 値の調査では、各分類項目のエダマメが 20 個写っているテスト画像を 10 反復分を使用した。したがって、誤分類率と未検出率は、誤分類数と未検出数をそれぞれ分類項目ごとに 200 で除して百分率で算出した。

4) 精度評価全般の補足について

精度評価用の画像は、重なり合っているエダマメや画面の端に見切れているエダマメが写っていないものを使用した。精度評価の正判定の基準は、次の 2 つである。

- ・ AI がエダマメを検出した際に表示する矩形（Bounding box）の中にエダマメさやが半分以上入っている場合
- ・ AI が正解の分類項目へ分類した場合

一方、誤判定の基準は、次の 3 つである。

- ・ テスト画像内のエダマメ以外の場所を範囲指定した場合
- ・ 品質分類を間違えた場合
- ・ エダマメを未検出した場合

なお、欠粒と一部の変色、さや割れの正判定は、アノテーションで部分的に範囲指定を行ったため、矩形内に傷害部分が半分以上入っている場合とした。また、ニュートン効率の調査では、1 つのエダマメに 2 つ以上の検出と分類が行われた場合を誤判定とした。AI の正判定と誤判定の一例を図 2-10 に示す。図 2-10 は、良品のエダマメが 4 個写っているテスト画像である。この例は誤分類の例であり、AI が右上と右下の良品をそれぞれ欠粒（ketsuryu）と奇形（kikei）へ誤分類している。

物体検出の設定に関し、信頼度は一般に設定される 50%とした。信頼度は、AI が領域推定した矩形内に各分類項目のいずれかが存在する確率値である。また、AI へ入力する画像サイズは、YOLOv3 の物体検出 AI で 768×768 ピクセル、Faster R-CNN の AI で 1,280×1,024 ピクセルとした。Faster R-CNN の精度評価では、用意した物体検出プログラムの仕様のため、テスト画像を MP4 形式の動画へ変換してから入力した。

本実験の統計処理は、分散分析を Excel 2013(Microsoft)で行い、多重比較を SAS 9.4(SAS Institute Japan)で行った。多重比較は Tukey 法で行い、各統計処理の有意水準は 5%とした。

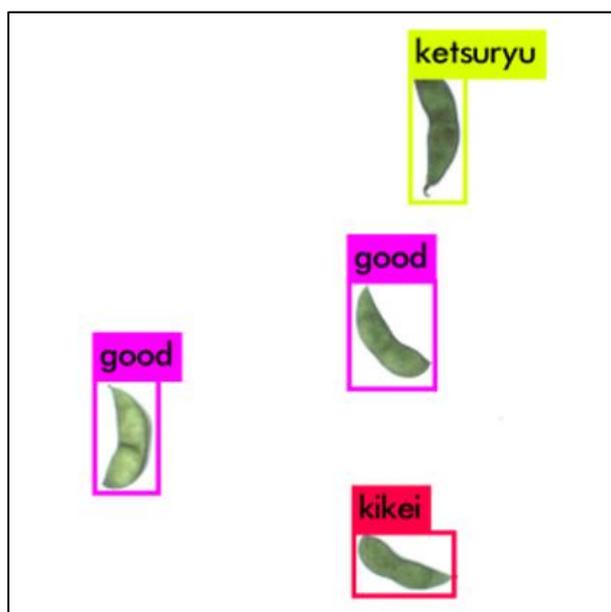


図 2-10 AI による良品エダマメの正判定と誤判定例

第3節 結果と考察

(1) 適合率・再現率・F値

適合率と再現率，F値の結果を表2-4～表2-9に示す。表の各値は，10反復のテストデータで算出した結果の平均値である。

YOLOv3に関し，各指標の平均は適合率がデータセットGのAIで0.91，再現率がデータセットEのAIで0.86，F値がデータセットEとGのAIで0.86とそれぞれ最高となった。一方，データセットCのAIは各指標の平均が全データセット中で最低となり，それぞれ0.73，0.67，0.67であった。

表2-4 データセットに含まれる品種の違いが各分類項目の適合率に与える影響 (YOLOv3)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.66 a	0.94 a	0.92 a	0.76 a	0.98 a	0.99 a	0.92 a	0.88 ab
B	0.52 bc	0.94 a	0.85 a	0.60 a	0.97 a	0.97 ab	0.88 a	0.82 c
C	0.45 c	0.91 a	0.69 b	0.70 a	0.48 b	0.92 c	0.93 a	0.73 d
D	0.68 a	0.94 a	0.92 a	0.70 a	0.99 a	0.99 a	0.91 a	0.88 ab
E	0.72 a	0.97 a	0.95 a	0.78 a	0.95 a	0.96 ab	0.97 a	0.90 ab
F	0.62 ab	0.94 a	0.87 a	0.70 a	0.97 a	0.96 ab	0.92 a	0.85 bc
G	0.73 a	0.95 a	0.96 a	0.78 a	0.99 a	0.97 ab	0.95 a	0.91 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

表 2-5 データセットに含まれる品種の違いが各分類項目の再現率に与える影響 (YOLOv3)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.89 ab	0.96 a	0.78 ab	0.89 abc	0.61 a	0.83 b	0.76 a	0.82 ab
B	0.82 bc	0.96 a	0.70 b	0.86 bc	0.20 c	0.86 ab	0.73 a	0.73 c
C	0.77 c	0.79 a	0.56 c	0.74 d	0.18 c	0.92 a	0.73 a	0.67 d
D	0.88 ab	0.96 a	0.82 ab	0.96 a	0.44 ab	0.87 ab	0.83 a	0.82 a
E	0.93 a	0.97 a	0.85 a	0.92 ab	0.60 a	0.94 a	0.82 a	0.86 a
F	0.87 ab	0.95 a	0.82 ab	0.81 cd	0.28 bc	0.92 a	0.82 a	0.78 bc
G	0.87 ab	0.99 a	0.85 a	0.94 ab	0.56 a	0.91 ab	0.79 a	0.84 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

表 2-6 データセットに含まれる品種の違いが各分類項目の F 値に与える影響 (YOLOv3)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.76 a	0.95 a	0.84 ab	0.82 ab	0.74 a	0.90 a	0.83 a	0.83 ab
B	0.64 bc	0.93 a	0.76 b	0.70 a	0.32 c	0.91 a	0.80 a	0.72 c
C	0.57 c	0.85 b	0.61 c	0.71 c	0.25 c	0.92 a	0.81 a	0.67 c
D	0.77 a	0.95 a	0.87 a	0.80 ab	0.60 ab	0.92 a	0.87 a	0.82 ab
E	0.81 a	0.97 a	0.89 a	0.84 a	0.72 a	0.95 a	0.89 a	0.86 a
F	0.72 ab	0.94 a	0.84 ab	0.75 bc	0.42 bc	0.94 a	0.86 a	0.78 b
G	0.79 a	0.97 a	0.90 a	0.85 a	0.70 a	0.94 a	0.86 a	0.86 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

Faster R-CNN に関し、各指標の平均はデータセット G の AI が最も高く、それぞれ 0.85, 0.83, 0.83 であった。一方、データセット B の AI は各指標の平均が全データセット中で最低となり、それぞれ 0.76, 0.73, 0.71 であった。YOLOv3 と Faster R-CNN の各データセットでの適合率と再現率、F 値の平均は AI 間で有意差があった。

表 2-7 データセットに含まれる品種の違いが各分類項目の適合率に与える影響 (Faster R-CNN)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.73 ab	0.91 a	0.97 a	0.63 cd	0.56 a	0.96 a	0.95 a	0.82 ab
B	0.53 c	0.91 a	0.92 a	0.55 d	0.58 a	0.95 a	0.87 bc	0.76 c
C	0.64 bc	0.90 a	0.88 a	0.53 d	0.57 a	0.95 a	0.90 abc	0.77 c
D	0.76 ab	0.93 a	0.96 a	0.69 bc	0.54 a	0.99 a	0.97 a	0.84 a
E	0.67 ab	0.95 a	0.90 a	0.74 ab	0.58 a	0.97 a	0.95 a	0.82 a
F	0.67 ab	0.92 a	0.89 a	0.68 bc	0.53 a	0.97 a	0.85 c	0.78 bc
G	0.76 a	0.95 a	0.92 a	0.82 a	0.59 a	0.98 a	0.94 ab	0.85 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

表 2-8 データセットに含まれる品種の違いが各分類項目の再現率に与える影響 (Faster R-CNN)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.69 a	0.96 a	0.52 cd	0.91 a	0.95 abc	0.84 a	0.79 a	0.81 ab
B	0.80 a	0.95 a	0.42 d	0.81 a	0.86 d	0.51 c	0.78 a	0.73 c
C	0.71 a	0.89 ab	0.58 bc	0.80 a	0.88 cd	0.64 bc	0.75 a	0.75 c
D	0.72 a	0.96 a	0.60 bc	0.86 a	0.96 ab	0.77 a	0.85 a	0.82 a
E	0.79 a	0.91 ab	0.71 ab	0.81 a	0.97 ab	0.71 ab	0.78 a	0.81 ab
F	0.69 a	0.86 b	0.64 abc	0.78 a	0.90 bcd	0.71 ab	0.79 a	0.76 bc
G	0.76 a	0.90 ab	0.77 a	0.81 a	0.97 a	0.81 a	0.86 a	0.83 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

表 2-9 データセットに含まれる品種の違いが各分類項目の F 値に与える影響 (Faster R-CNN)

データセット名	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	平均
A	0.70 ab	0.93 a	0.67 cd	0.74 ab	0.71 a	0.89 a	0.86 ab	0.79 bc
B	0.64 b	0.93 a	0.56 d	0.65 bc	0.69 a	0.66 b	0.82 ab	0.71 e
C	0.67 ab	0.89 a	0.69 bc	0.63 c	0.69 a	0.75 bc	0.81 b	0.73 de
D	0.73 ab	0.94 a	0.74 abc	0.76 a	0.69 a	0.86 a	0.90 a	0.80 ab
E	0.72 ab	0.93 a	0.79 ab	0.77 a	0.72 a	0.81 ab	0.85 ab	0.80 abc
F	0.70 ab	0.89 a	0.74 abc	0.72 abc	0.66 a	0.81 ab	0.82 b	0.76 cd
G	0.76 a	0.92 a	0.83 a	0.80 a	0.73 a	0.88 a	0.89 ab	0.83 a

注：AI間において異なるアルファベット間に有意差あり（有意水準5%）

適合率と再現率，F 値の平均は，データセット A, B, C (1 品種の画像のみ含む) の AI を比較した場合，YOLOv3 と Faster R-CNN の両方でデータセット A の AI が有意に高かった。この精度変化は，各品種の外観の違いに起因すると考える。庄内三号とグリーン75の違いは毛茸の色であり，それぞれ茶色と白色である⁶⁵⁾。庄内三号と庄内一号は山形県庄内地方特産のダダチャマメ系のエダマメであり，外観の特徴はほぼ同じである⁶⁶⁾。しかし，庄内一号は百粒重が庄内三号よりも重く，さやが一回り大きい。このように，エダマメの外観は品種ごとに差異があり，同系統の品種間にもわずかな差異がある。AI の精度は，DL が各分類項目の特徴を正確に学習しているのであれば学習データの品種の違いによって変化しないと考ええる。ところが，本実験では，学習データの品種の変化で精度に差が生じた。よって，エダマメの精選別用 AI を DL で開発する場合，各品種の外観の差異は AI の精選別精度に影響を与えるといえる。そして，AI は選別対象と異なる品種の画像のみで開発すると，精度が低下するといえる。

次に，適合率と再現率，F 値の平均は，データセット D, E, F, G (複数の品種の画像を含む) の AI を比較した場合，YOLOv3 と Faster R-CNN の両方でデータセット F の AI が低くなった。この原因は，データセット F に選別対象の庄内三号の画像が含まれていないためと考える。以上から，高精度なエダマメの精選別用 AI の開発には，選別対象品種の画像を必ず学習データに含める必要があるといえる。

また，データセット D, E, G の AI は，データセット A の AI と同等の精度となった。これは，選別対象の品種のデータ含有率が低い場合でも，AI の精度が選別対象品種のデータのみで開発した場合と同程度になることを示している。そのため，開発者は生産量の少なさなどの理由で選別対象品種の画像を十分に収集できなかった場合，他の品種の画像を学習データに混合しても高精度な AI を開発できるといえる。これらの知見は，

エダマメ精選別用の学習データのルール決定に役立つ。

(2) ニュートン効率

ニュートン効率の結果を図 2-11~2-12 に示す。図のエラーバーは標準偏差を、アルファベットは多重比較の結果を示す。

YOLOv3 では、データセット D の AI が 0.79 と最高値を示した。そして、Faster R-CNN では、データセット E の AI が 0.68 と最高値を示した。YOLOv3 の AI では、データセット C とその他のデータセットの AI 間でニュートン効率に有意差が見られた。

片平ら¹⁵⁾の研究では、手選別のニュートン効率が 0.59 であることが報告されている。人間のニュートン効率が最大値の 1 ではない理由は、人間も分類を間違えることがあるためである。YOLOv3 のデータセット D の AI はニュートン効率が 0.79 であり、本研究で開発した AI の中で最も精度が高かった。よって、DL で開発した AI は、良品と不良品の 2 項目の選別精度が手選別よりも優れているといえる。

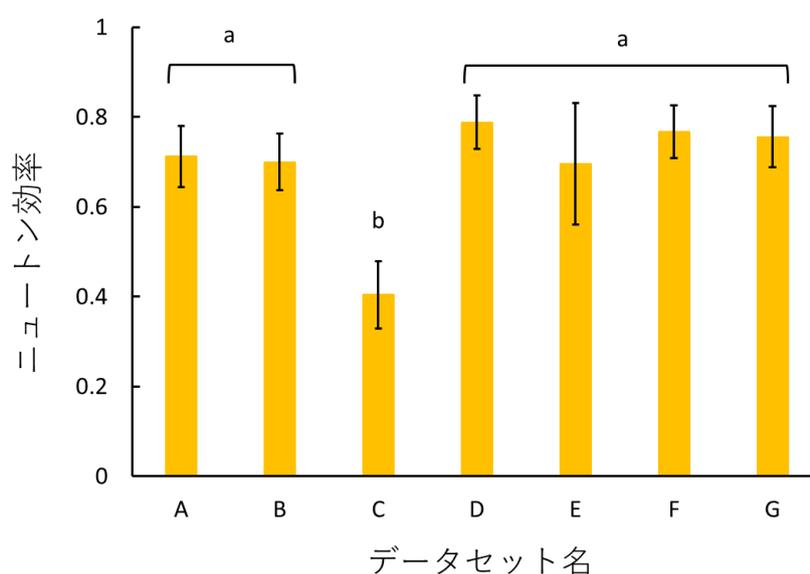


図 2-11 データセットに含まれる品種の違いがニュートン効率に与える影響 (YOLOv3)

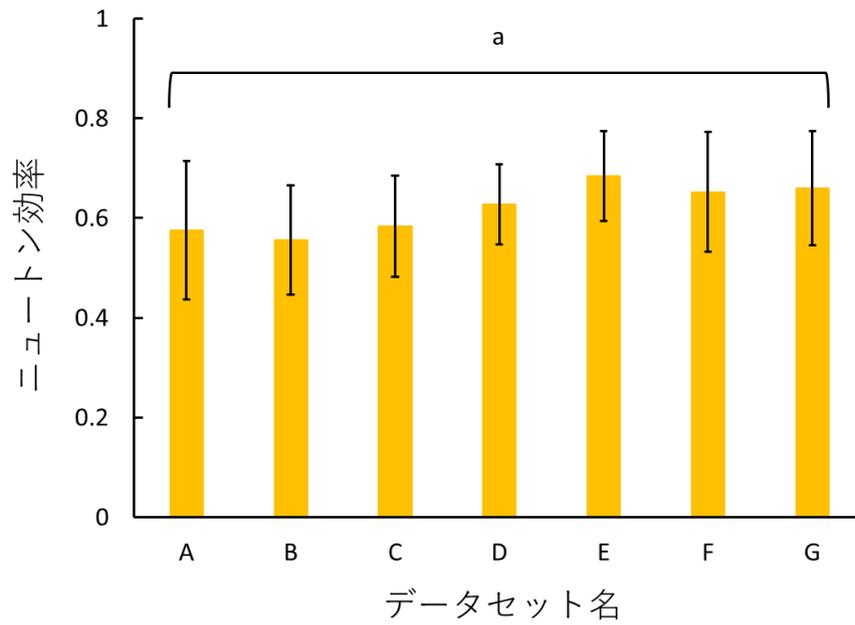


図 2-12 データセットに含まれる品種の違いがニュートン効率に与える影響 (Faster R-CNN)

(3) 誤分類率と未検出率

YOLOv3 と Faster R-CNN で最も精度が優れていた AI は、データセット G と判断した。データセット G の AI は F 値の平均が最高で、かつ高いニュートン効率を示したためである。データセット G の AI の誤分類率と未検出率の結果を表 2-10~2-11 に示す。

表 2-10 データセット G で開発した YOLOv3 の誤分類率と未検出率

テスト画像の 分類項目	誤分類率							未検出率
	AIが誤分類したときの分類結果							
	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	
良品		0	0	9	0	0	2	5
一粒	0		1	1	0	0	1	0
欠粒	5	0		2	1	1	1	8
奇形	3	1	0		0	2	1	3
未熟	24	4	2	11		0	0	7
変色	0	0	0	4	0		1	6
さや割れ	1	1	1	3	0	2		16

注：単位は「%」

表 2-11 データセット G で開発した Faster R-CNN の誤分類率と未検出率

テスト画像の 分類項目	誤分類率							未検出率
	AIが誤分類したときの分類結果							
	良品	一粒	欠粒	奇形	未熟	変色	さや割れ	
良品		0	3	3	23	0	0	0
一粒	0		0	0	11	0	0	0
欠粒	8	1		4	18	1	1	0
奇形	8	1	1		9	2	2	1
未熟	0	3	1	2		0	0	0
変色	7	0	1	5	5		3	3
さや割れ	3	1	2	7	5	0		1

注：単位は「%」

1) 誤分類率

誤分類が多かった分類項目の上位 3 項目を表 2-12 に示す。YOLOv3 の AI は、「未熟」を「良品」へ誤分類した割合が 24%で最高となり、次いで「未熟」を「奇形」へ誤分類した割合が 11%、「良品」を「奇形」へ誤分類した割合が 9%であった。次に、Faster R-CNN の AI は、「良品」を「未熟」へ誤分類した割合が 23%で最高となり、次いで「欠粒」を「未熟」へ誤分類した割合が 18%、「一粒」を「未熟」へ誤分類した割合が 11%であった。

表 2-12 誤分類が多かった分類項目とその内容

	YOLOv3			Faster R-CNN		
	正	誤	誤分類率 (%)	正	誤	誤分類率 (%)
1	未熟	良品	24	良品	未熟	23
2	未熟	奇形	11	欠粒	未熟	18
3	良品	奇形	9	一粒	未熟	11

良品と未熟、そして欠粒の誤分類の原因は、これらの選別基準がエダマメ子実部のさやの厚みであるためと考える。良品と未熟は、すべての子実部のさやの厚みがそれぞれ 8mm 以上と 8mm 未満の個体である。そして、欠粒は、子実部のさやの厚みが 8mm 以上と 8mm 未満の部分が混在する個体である。本実験の AI は、エダマメを真上から撮影して高さ情報を含まない 2 次元の俯瞰画像を使用して、開発と精度評価が行われた。人間は、これら 3 項目の選別を行う場合、外観と一緒にさやの厚みを確認する。一方、AI はさやの具体的な厚みを計測しておらず、俯瞰画像の画素情報のみを利用して特徴抽出や検出・分類を行った。AI は俯瞰画像から分類に必要な特徴量を抽出できない場合に誤分類を起こす。したがって、開発者は、厚みの計測値を基準に選別する分類項目に関して、AI が具体的な計測値を測って検出・分類していないことを情報共有する必要がある

る。

AIにこれら3項目をより高精度に検出・分類させるには、AIが画素情報から特徴を抽出できる撮影環境を整える必要がある。これら3項目の外観の違いは、子実部周辺に発生する影の濃淡である(図2-13)。良品は、真上から光を当てると子実と子実の間に影ができ、さらに子実の輪郭が影により強調される。未熟は子実の厚さが良品よりも小さいため、全体的に影の濃淡が薄い。そのため、未熟は、良品よりも子実の輪郭を明瞭に撮影されない。欠粒は、良品子実部と未熟子実部の境目を影の濃淡で区別できる。エダマメの精選別用AIでは、影の濃淡の特徴を明瞭に撮影できるように注意することで精度を改善できるといえる。そのため、開発者は撮影環境やカメラのキャリブレーション方法を統一し、データ収集のルールにそれらを加える必要がある。



図2-13 真上から撮影したグリーン75の良品と未熟、欠粒の外観の差異

また、AIの誤分類は、先述の項目以外でも発生した。この原因は、1つのエダマメに2項目以上の要素が混在することがあるためと考えられる。本実験で使用したエダマメは、1つにつき2つ以上の分類要素を持つ個体があった。それらのエダマメは、手選別時に選別者の主観で7項目の分

類項目のうちいずれかへ分類された。その後、それらは、アノテーションで1つの分類ラベルを付与された。AIは、DLでアノテーション情報と同様の検出・分類を行えるようにネットワークが調整される。このように開発されたAIは、ラベル付けされなかった分類要素の画素情報が入力された場合、それを検出・分類しないようになる。さや割れと変色の2つの要素を持つエダマメに変色の正解ラベルを付けた場合では（図 2-14）、AIが今後同じようなさや割れを検出・分類しないようになり、精度の低下に繋がる（図 2-15）。

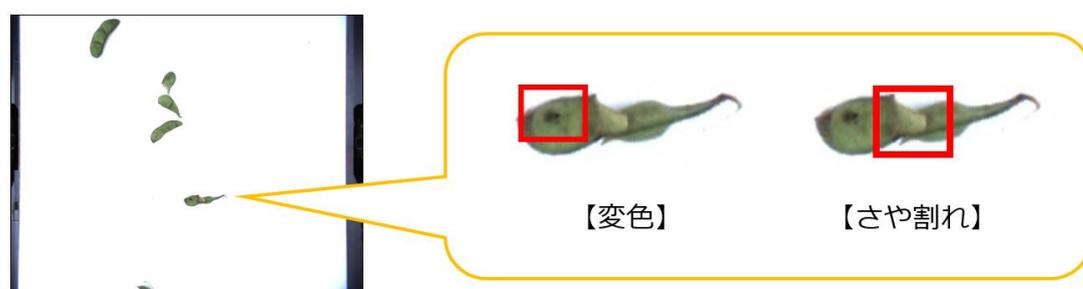


図 2-14 さや割れと変色の2つの要素を持つエダマメの例

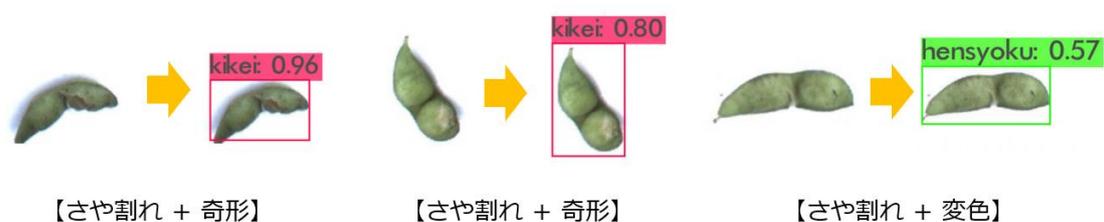


図 2-15 データセット G で開発した YOLOv3 の物体検出 AI が
さや割れのエダマメを誤分類した例

複数の分類要素を含むエダマメが写っていた場合の対処方法は、分類項目の優先順位を事前に決めてアノテーションすることが考えられる。

これにより、AI が分類の優先順位を学習し、複数の分類要素が含まれる場合に優先順位の高い方へ選別するようになると思う。分類項目の優先順位は、規格外品を最優先とし、次いで B 品、A 品を設定するとよいと思う。さらに、規格外品や B 品はそれぞれ複数の項目があり、それらの中でも優先順位を決める。これは、実際の選別では規格外品を最優先で取り除く必要があり、AI に規格外品のエダマメを確実に検出・分類させるためである。エダマメの出荷規格は各地域で公布されている⁶⁷⁾。開発者は、エダマメの精選別用 AI の開発に際して各地の出荷規格を整理し、選別の優先順位をアノテーションの前に明確化するとよい。

2) 未検出率

未検出率は、YOLOv3 で「さや割れ」が、Faster R-CNN で「変色」が最も高かった。この 2 項目の未検出の原因は、傷害がさやの一部に発生する可能性があるためと考えられる。この 2 項目は、アノテーションの範囲指定をさやの傷害部分に限定した。さやの一部にアノテーションの範囲を限定した場合、AI はアノテーションの指定範囲通りに検出できるように調整される。そのため、AI がさやの一部にさや割れや変色の特徴を抽出した場合、AI はさや割れや変色を検出・分類できる。しかし、CNN はさやの中に他の分類要素が存在する場合、それらの特徴も抽出する。AI による特徴量をベースとした分類確率の算出では、さや割れや変色の分類確率が信頼度閾値より低い場合、クラス分類確率が高い他の分類項目に分類される。また、AI は他の分類項目の分類確率が信頼度閾値を下回った場合、画像中に検出対象物が無いとして未検出判定をする。さや割れや変色のエダマメは、他の分類項目の要素も含んでいることが多いため、アノテーション範囲をさやの一部に限定した分類項目は、さやの全体を範囲指定した分類項目よりも未検出率が上昇する。

この問題の解決策は、アノテーションでさや全体を範囲指定することである。これにより、AIはさや割れや変色を部分的に検出しなくなり、1つのエダマメにつき1つの分類を行うようになる可能性がある。

3) AI のネットワーク構造の違いによる誤分類と未検出の内容の変化について

誤分類と未検出が多い分類項目は、YOLOv3 と Faster R-CNN で異なっていた。これは、AI の Head 部や CNN の違いによるものであると考える。AI に入力された画像は、最初に CNN により特徴マップへ変換される。CNN は、種類によって畳み込み層やプーリング層などの構成が異なる。そのため、CNN 構造の違いは、抽出される特徴量の変化につながる。さらに、YOLOv3 と Faster R-CNN は、特徴マップからの検出と分類の仕組みが異なる。本研究で開発した YOLOv3 と Faster R-CNN の AI は、これらの理由から誤分類と未検出の多い項目が異なつたと考える。開発者は、AI の Head 部と CNN の種類の違いによって、品質分類の誤分類と未検出の内容が変化することを考慮すべきである。

第4節 摘要

- 1) 適合率と再現率，F 値の平均は，選別対象品種の画像を学習データに含む AI の方が，他品種の画像のみを学習データに含む AI よりも有意に高くなった。
- 2) AI の精度は，選別対象品種とその他の品種の混合割合が変化しても低下しなかったため，選別対象品種のデータ不足を他品種のデータで補うことができる。
- 3) YOLOv3 と Faster R-CNN の F 値の平均は，いずれも学習データに選別対象の品種が含まれていないと有意に低下したため，品種の組み合わせが様々な構造の AI の総合的な精度に共通の精度変化をもたらす。
- 4) 本研究で開発した YOLOv3 の AI は，ニュートン効率が最大で 0.79 となり，手選別のニュートン効率を 0.20 上回り，良品と不良品の 2 項目の選別精度が手選別よりも優れていた。
- 5) 誤分類は YOLOv3 で未熟，Faster R-CNN で良品が多かった。未検出は YOLOv3 でさや割れ，Faster R-CNN で変色が多かった。AI の誤分類と未検出は，AI の Head 部と CNN の違いで変化した。
- 6) 誤分類は未熟度に関する良品と未熟，欠粒の 3 項目間で多発し，子実部周辺の影の濃淡が影響したことから，影の濃淡を鮮明にして撮影する必要がある。
- 7) 未検出はアノテーション範囲をさやの一部に限定した変色，さや割れで多発した。これは，1 つのさやに 2 つ以上の分類要素が含まれている場合にクラス分類確率の低下で引き起こされ，すべてのエダマメでさや全体をアノテーションすることで解決できる可能性がある。

第Ⅲ章

CNN の分類根拠可視化技術を利用した

AI の精度変化要因の解明

第 1 節 緒言

(1) 第Ⅱ章の実験結果と AI の精度変化の要因

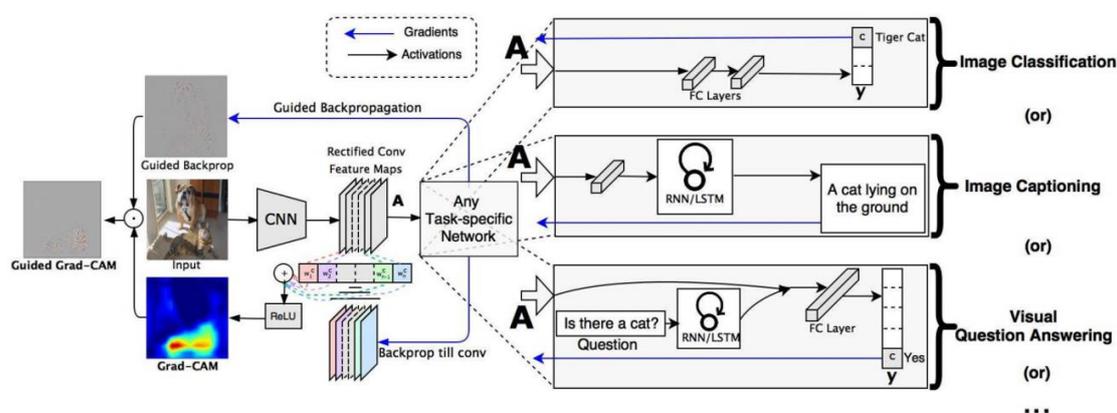
第Ⅱ章の実験から、学習データに含む品種の違いが物体検出 AI の精度変化に影響することに加え、学習データに選別対象品種のデータを含めることで物体検出 AI の精度が向上することが分かった。しかし、第Ⅱ章の実験において、すべての学習データセットは、分類項目の選別基準を統一していた。AI の精度は、DL で各分類項目の外観の特徴のみを正確に抽出できるように学習していれば、品種に関わらず一定になると考えられる。よって、AI は DL で品種固有の外観の特徴も一緒に学習した可能性がある。しかし、現時点では、品種間の外観の差異が AI の分類精度に変化をもたらす要因であるかが断定できていないため、AI の分類根拠を理解できれば、AI の特性に基づいた学習データを構築できる可能性がある。

(2) 説明可能な AI による分類根拠の特定

DL は、学習データに基づいて AI のネットワークを自動調整する。DL で開発された AI は、学習データのどの情報を最終的な検出・分類へ利用したか不明瞭となりブラックボックス化するため、入力データ内の対象物と関係のない部分を検出や分類に利用している可能性もある。これは、将来新しいデータを入力した際に、誤検出や誤分類の原因となる。

近年、DL の利用増加に伴い、「説明可能な AI (XAI: Explainable Artificial Intelligence)」⁶⁸⁾が求められるようになった。説明可能な AI は、入力データに対する AI の分類根拠を明確化して、利用者からの信頼性を担保する

ための技術である。2017年には、「Grad-CAM」⁶⁹⁾が CNN の分類根拠可視化技術として登場した。Grad-CAM は、分類スコアへの影響が大きい部位を勾配（微分係数）の平均化によって特定する技術である。勾配が大きい箇所は、分類に大きく寄与した部分として、ヒートマップで強調される。開発者は、そのヒートマップを確認することで対象物の分類に重要な部位を特定できる（図 3-1）。

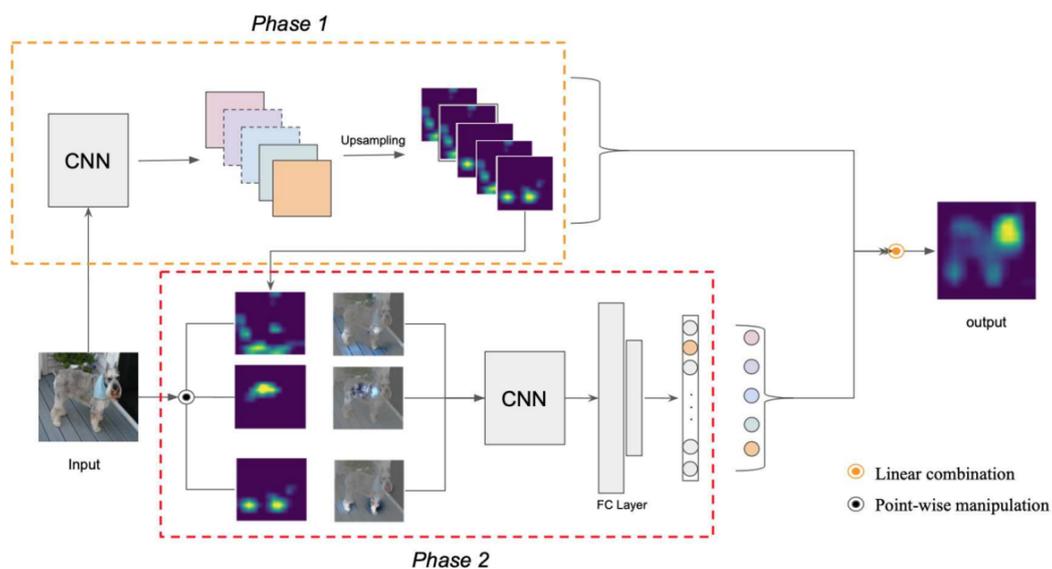


引用：Ramprasaath, S., Michael, C., Abhishek, D., Ramakrishna, V., Devi, P., Dhruv, B., 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, IEEE International Conference on Computer Vision (ICCV), 618-626. Figure 2.

図 3-1 Grad-CAM による分類根拠の可視化までの流れ

そして、2020年に「Score-CAM」⁷⁰⁾が登場した。Score-CAMは、Grad-CAMのように勾配を利用せず、順伝播で重みづけを行う分類根拠可視化技術である。Grad-CAMなどは、クラス分類時の勾配を利用することで勾配消失の影響を受けやすくなる。そして、分類根拠可視化精度は分類に重要な重みにノイズが混入することで、低下することがある。そのため、Score-CAMは順伝播で重みづけを行い、分類根拠可視化精度を向上させた。Score-CAMは、最初に入力データをCNNへ入力して特徴マップを得る。次に、特徴マップのサイズを入力データと同じサイズへ拡大し、特徴マップの各値を0から1の間になるように正規化する。そして、入力データと正規化した特徴マップを組み合わせたマスク画像を再度CNNで

分類し、対象の分類クラスのスコアを取得する。その後、このマスク画像に対する分類スコアと元の入力データの分類スコアを引いた値に softmax 関数を施して、重みを算出して足し合わせる。この操作は、全チャンネルで行い、最終的に分類上重要とされる部位をヒートマップで出力する（図 3-2）。



引用：Haofan, W., Zifan, W., Mengman, D., Fan, Y., Zijian, Z., Sirui, D., Piotr, M., Xia, H., 2020. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 24-25. Figure 3.

図 3-2 Score-CAM による分類根拠の可視化までの流れ

Grad-CAM や Score-CAM は、DL で開発した AI の分類根拠を解明するために、多数の事例で応用されている。西尾ら⁷¹⁾は、橋梁の損傷度を 3 段階で判定する画像認識 AI を開発した。そして、西尾らは、Grad-CAM を利用して、画像認識 AI が入力データの損傷部を分類に重要としているか確認した。この研究では、画像認識 AI が正常にクラス分類を行えた場合であっても、損傷部位に注目していない場合があった。この結果を基に、西尾らは、画像認識 AI が実際の損傷部位を重要視していることを確認した画像を利用して画像認識 AI を再学習したことで、高精度な画像認

識 AI を開発できたことを報告している。次に、Yebasse ら⁷²⁾はコーヒーの葉の病気を画像認識 AI で分類し、Grad-CAM や Score-CAM で分類根拠を可視化した。この研究では、画像認識 AI が一部の画像で背景の土を病気の特徴としていることが分かった。この結果から、Yebasse らは、画像認識 AI の分類精度向上のために、U2-Net⁷³⁾を利用して画像中の葉の領域のみを切り出す前処理を行った。そして、葉の病気の分類精度はそれらの切り抜き画像を使用することで向上するといった手法を提案した。また、Yebasse らは、Score-CAM が Grad-CAM よりもより緻密に分類根拠を特定できることを確認した。以上から、開発者は CNN の分類根拠可視化技術で AI が入力データのどの部位を分類上重要としていたか確認し、その結果を基に学習データを再構成することで高精度な AI を開発できるといえる。

(3) 本実験の目標

AI の分類根拠が学習データ内の特定の条件の違いで変化した場合、その条件が AI の精度変化の因子であることを断定できる。そして、開発者は精選別用の AI の特性を理解し、その因子の影響を考慮した学習データを構築できるようになる。

本実験の目標は、Score-CAM を使用し、学習データの品種の違いが AI の精度変化の要因であるか特定することである。そして、その結果を基に、品種の影響を受けない学習データの構築方法を考案する。

第 2 節 実験方法

本実験のフローは、図 3-3 の通りである。最初に、3 品種のエダマメをそれぞれ 7 通りに手選別し、各品種の画像を収集した。次に、撮影画像からエダマメを 1 さやずつ切り抜き保存し、品種ごとに画像認識 AI（以下、本章では「AI」と称する）開発用及び AI の精度評価・Score-CAM 用の 2 つへ振り分けた。その後、品種の異なる 3 つの学習データセットで DL を行い、AI を開発した。最後に、精度評価用の画像を各 AI へ入力し、AI の精度評価と Score-CAM による分類根拠の特定を行った。

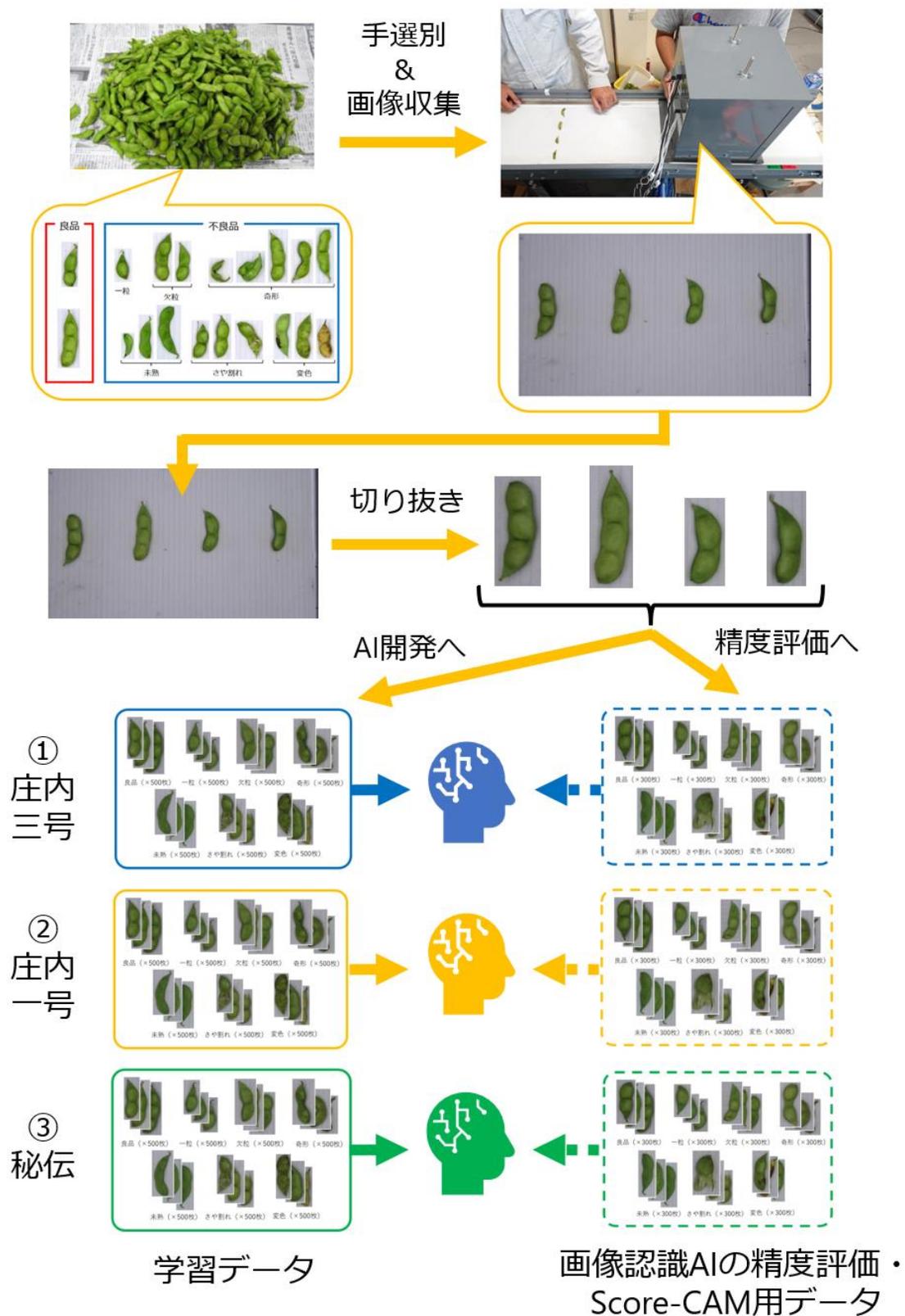


図 3-3 品種ごとの画像認識 AI の開発から分類根拠可視化までの流れ

(1) 供試材料と画像の収集

本実験で使用したエダマメは、2020年に山形大学農学部附属やまがたフィールド科学センター高坂農場で栽培した。品種は、「庄内三号」と「庄内一号」、「秘伝」の3品種である。秘伝は、白毛の大きさやで、晩生の品種である（表3-1）。本実験でこれらの品種を使用した理由は、AIの分類根拠が各品種の外観の差異によりどのように変化するか確認するためである。一般に、極早生品種と早生品種、中生品種は、開花から収穫までに30～40日を要する。一方、晩生品種は開花から収穫までに40～50日を要する⁶⁵⁾。そのため、晩生品種は他品種よりもさやと子実の形成期間が長く、さやの形状や色合いが異なる特徴がある。これら3品種を使用することで、同系統で外観の特徴がほとんど一致する品種と、早晩性の違いで外観の特徴が異なる品種の分類根拠を比較できると考える。選別者は、エダマメの収穫後に、品種ごとに良品1項目と不良品6項目の計7通りに手選別した（図3-4）。本実験の手選別の基準は、第2章と同様である。2つ以上の分類要素を持つエダマメは、選別者の主観で特徴の大きい方に分類した。

表 3-1 各エダマメ品種の特徴

品種名	子実の色	早晩性	さやの毛色	外観の特徴
庄内三号	茶	中生	茶	さやの子実間のくびれ
庄内一号	茶	早生	茶	さやの子実間のくびれ
グリーン75	青	極早生	白	大さや
秘伝	青	晩生	白	大さや

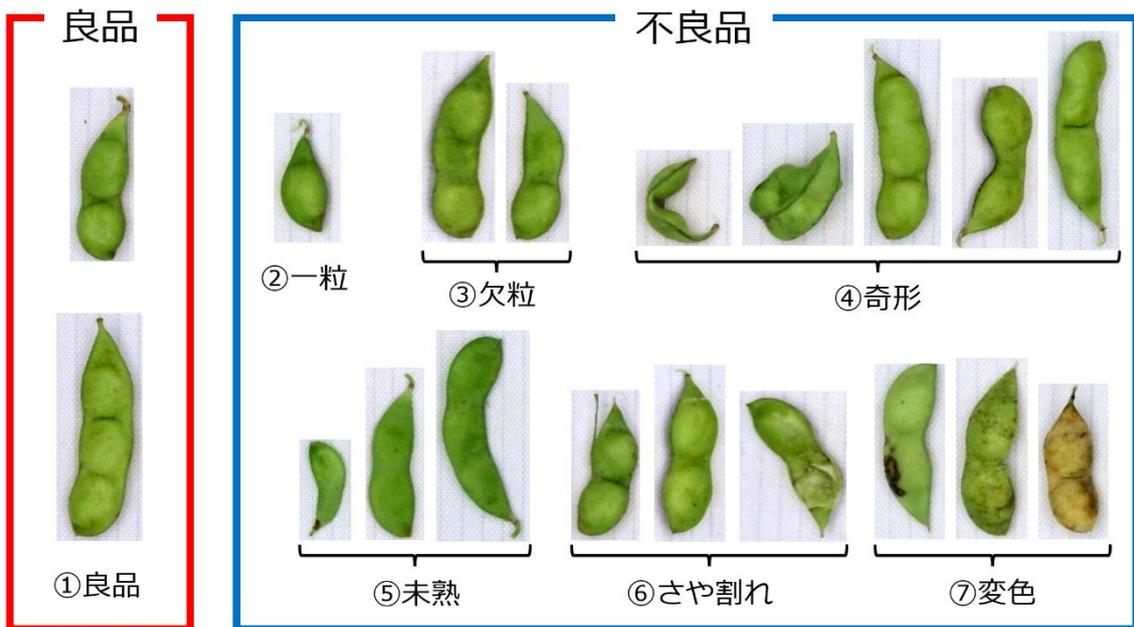


図 3-4 7 項目のエダマメ分類項目

画像の収集は、図 3-5 の撮影機を使用して行った。この撮影機は、白色の搬送ベルトと、立方体の撮影部で構成される。撮影部の内部は、天井に 6 本の LED バーライトと、1 個の Web カメラ (KAYETON, OV4689) を備えている。Web カメラからエダマメまでの距離は約 310mm であり、撮影の縦幅が 142mm、横幅が 256mm である。画像のフォーマットは JPEG であり、解像度は 1,920×1,080 ピクセルである。

エダマメは品種ごとに撮影し、さらに分類項目ごとに撮影した。庄内三号と庄内一号は一度に 4 さやずつ撮影し、一方、秘伝は一度に 3 さやずつ撮影した。これは、秘伝を 4 さや並べると、画面の外にさやがはみ出すことがあったためである。エダマメは、1 枚撮影する度にさやの中心を軸に 45 度ずつ傾けて撮影した (図 3-6)。さらに、1 周回転させたエダマメは、その後裏返し、同じように 45 度ずつ傾けて撮影した。この撮影方法は、短時間で様々な角度のエダマメを撮影するために行った。

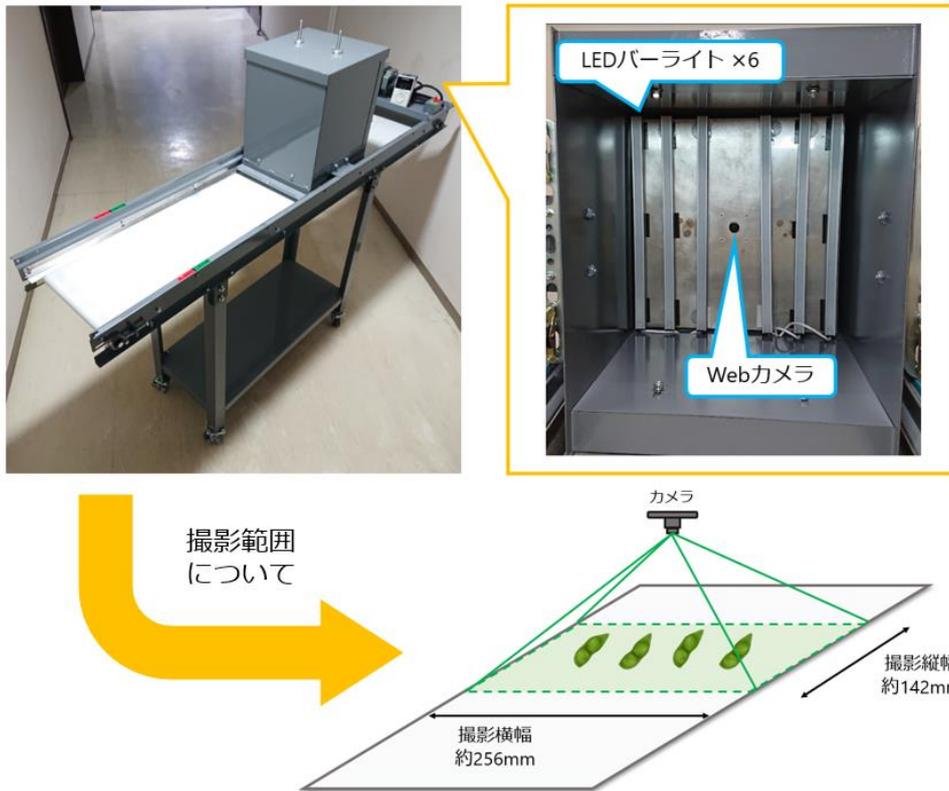


図 3-5 本実験で使用したエダマメ撮影機

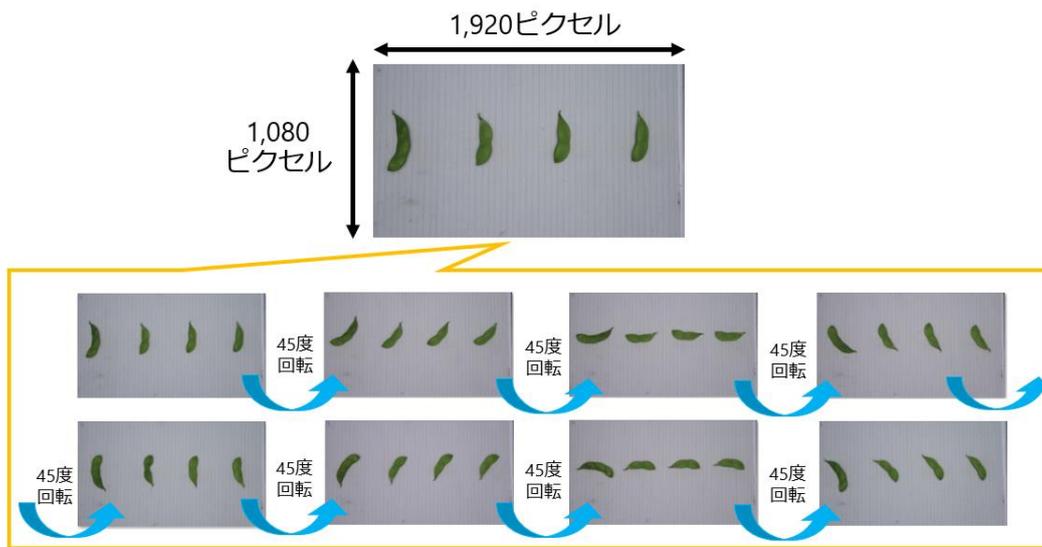


図 3-6 本実験で収集した 45 度ずつ回転させて撮影したエダマメ画像例

撮影機でエダマメ画像を収集した後、各画像からエダマメを 1 さやずつ切り抜いた。エダマメの切り抜きには、Microsoft 社のフリーソフトウェア「VoTT-2.2.0」と、Python3 のプログラムで行った (図 3-7)。最初に、VoTT-2.2.0 で画像内のエダマメを 1 個ずつ囲み、その座標情報を記録した Json ファイルを出力した。次に、Json ファイルの情報を元に、Python3 のプログラムでエダマメを切り抜き保存した。これらの切り抜き画像は、AI の開発と精度評価、Score-CAM に使用した。

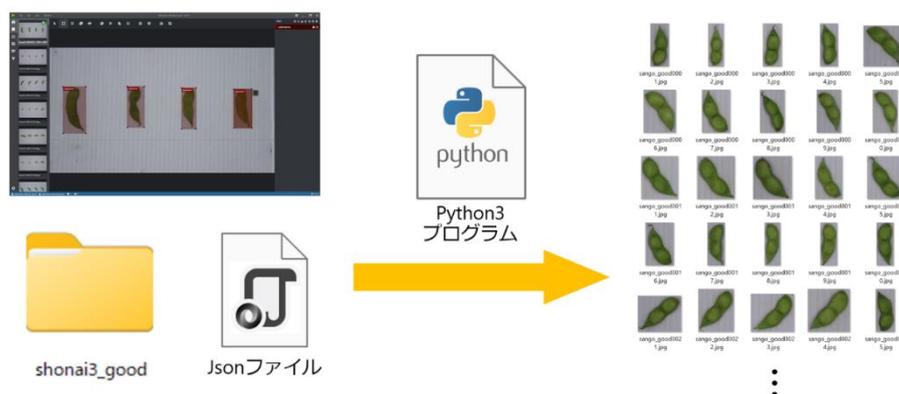


図 3-7 撮影画像からエダマメを 1 個ずつ切り抜き保存するまでの流れ

(2) 学習データと精度評価用のデータの準備

学習データ用と AI の精度評価用の画像数を表 3-2 に示す。学習データ用の画像は、各品種で分類項目毎に 500 枚用意した。そして、AI の精度評価用の画像は、各品種で分類項目毎に 300 枚用意した。学習データ用と精度評価用の画像は、重複しないように分別した。本実験では、第 II 章と異なり、品種ごとに AI を開発した。学習データセットは、庄内三号、庄内一号、秘伝の 3 つである。

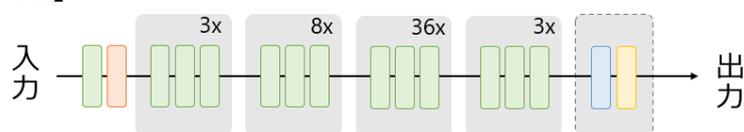
表 3-2 学習データと精度評価用の画像数

分類項目	学習データ数 (枚)			AIの精度評価用データ数 (枚)		
	庄内三号	庄内一号	秘伝	庄内三号	庄内一号	秘伝
良品	500	500	500	300	300	300
一粒	500	500	500	300	300	300
欠粒	500	500	500	300	300	300
奇形	500	500	500	300	300	300
未熟	500	500	500	300	300	300
さや割れ	500	500	500	300	300	300
変色	500	500	500	300	300	300
合計	3500	3500	3500	2100	2100	2100

(3) DL による AI の開発

本実験では、庄内三号と庄内一号、そして秘伝の 3 種類の学習データセットで ResNet152 と VGG16 の 2 種類の AI (CNN) を開発した (図 3-8)。2 種類の AI を開発した理由は、CNN の種類が異なる場合に、分類根拠がどのように変化するか調査するためである。ResNet152 は、152 層からなる CNN である。ResNet152 の特徴は、スキップ層とバッチ正規化を導入して、CNN の深層化に伴う特徴表現力の向上と勾配消失問題の解決を同時に達成したことである。また、ResNet152 は、2015 年の ILSVRC で最高の画像分類精度を達成し、さらに、人間の画像分類精度を超えたモデルである。VGG16 は、16 層からなる CNN である。VGG16 の特徴は、構造の簡素さと画像分類精度の高さであり、様々な用途で使用されてきた CNN である。

【ResNet152】



【VGG16】

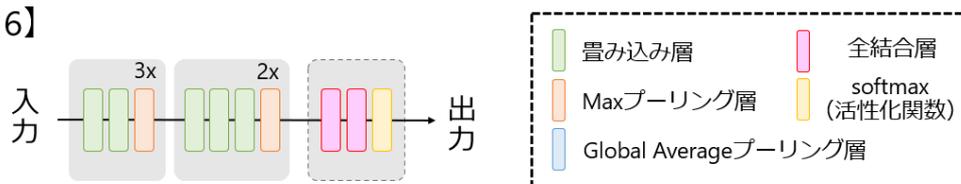


図 3-8 ResNet152 と VGG16 のネットワークの模式図

AI 開発用のフレームワークは、Tensorflow-gpu-1.15 (以下, Tensorflow) と Keras-2.3.1 (以下, Keras) を使用した。DL では、ImageNet による事前学習済みモデルをベースに転移学習を行った。CNN は上位層で複雑な特徴を、下位層で単純な特徴を抽出する⁷⁴⁾。そのため、AI は転移学習で上位層のパラメータを更新することで、より高精度化できると考える。したがって、ResNet152 では最終層から 32 層前までのパラメータを、VGG16 では最終層から第 15 層までのパラメータを固定し、それぞれ DL を行った。

DL 時のハイパーパラメーターに関し、DL の学習回数は、ResNet152 の AI ではすべて 10 回に設定した。一方、VGG16 の AI では庄内三号と庄内一号の学習データセットで 15 回、秘伝の学習データセットで 20 回に設定した。DL の回数がそれぞれ異なる理由は、過学習が発生する直前で DL を停止したためである。過学習の発生の確認は、Keras で DL の Train-Loss と Train-Accuracy, Validation-Loss, Validation-Accuracy を数値化できる機能を利用して行った。DL の停止タイミングは、これらの値が横ばいに推移し始めたときである。他のハイパーパラメーターは、学習データの Train と Validation の分割割合を 8:2、学習時の画像の入力サイズを 256×256 ピクセル、Batch サイズを 16 に設定した。本実験で使用したコンピューターの環境を表 3-3 に示す。

表 3-3 本実験で使用したコンピューターの環境

OS	Ubuntu20.04 LTS
CPU	Intel Core-i7 7700K
GPU	NVIDIA GeForce GTX1080Ti
NVIDIAドライバー	470
CUDA	10.0
cuDNN	7.6.5
OpenCV	4.5.0
Python	3.6

(4) AI の精度評価

本実験では、Score-CAM で各 AI の分類根拠を特定する前に、AI の精度評価を行った。各 AI の精度は、混同行列 (Confusion matrix) と Accuracy (正解率) で評価した。混同行列は正と負の二値分類の結果をまとめた表であり、AI が画像内のエダマメをどのように分類したか評価する指標である。混同行列は縦軸が精度評価用画像の正解の分類項目、横軸が AI の分類結果を示す (図 3-9)。

True label (正解の分類項目)	good	T	F	F	F	F	F	
	ichiryu	F	T	F	F	F	F	
	ketsuryu	F	F	T	F	F	F	
	kikei	F	F	F	T	F	F	
	mizyuku	F	F	F	F	T	F	
	sayaware	F	F	F	F	F	T	
	hensyoku	F	F	F	F	F	T	
			good	ichiryu	ketsuryu	kikei	mizyuku	sayaware
		Predicted label (AIによる分類結果)						

T : 外観品質を正しく分類した場合

F : 外観品質を誤って分類した場合

図 3-9 混同行列の模式図と各マスの意味

そして、Accuracy は、式 (5) で算出した。

$$\text{Accuracy (正解率)} = \frac{\text{AIが正しく外観品質を分類した数}}{\text{精度評価用の画像の総数 (=2100枚)}} \quad (5)$$

Accuracy は、AI がエダマメをどれだけ正確に分類できたか総合的に評価する指標である。

AI の精度評価では、各 AI の学習データセットに含む品種と同じ品種の画像を使用して行った。精度評価用の画像数は、分類項目ごとに 300 枚用意した (表 3-2)。

(5) Score-CAM による分類根拠の特定方法について

Score-CAM による分類根拠の特定は、図 3-10 の手順で行った。最初に、各品種の精度評価用画像の中から細分類項目ごとに 30 枚の画像を抽出した。細分類項目とは、7 項目の分類項目をさらに 19 項目へ細分化した項目である (図 3-11)。細分類項目ごとに分類根拠を可視化した理由は、同じ分類項目でも細分類項目ごとに分類根拠が変化している可能性を考慮したためである。次に、各 AI へ学習用データセットと同じ品種のテストデータを入力し、エダマメの分類と Score-CAM の適用を同時に行った。そして、AI がエダマメを正しく分類できた場合、Score-CAM のヒートマップの発生パターンを記録した。各 AI の分類根拠は、このヒートマップの発生パターンが最多のものとした。図 3-10 の例では、庄内一号の学習データセットで開発した AI が、さや割れの「割れ」のテストデータを分類している。そして、この AI が正しくさや割れに分類できた場合、花柱

を除くさや全体（花柱以外）にヒートマップが発生することが最多であった。よって、庄内一号の AI の「割れ」の分類根拠は、花柱を除くさや全体へ決定した。エダマメさやの各部位の名称を図 3-12 に示す。

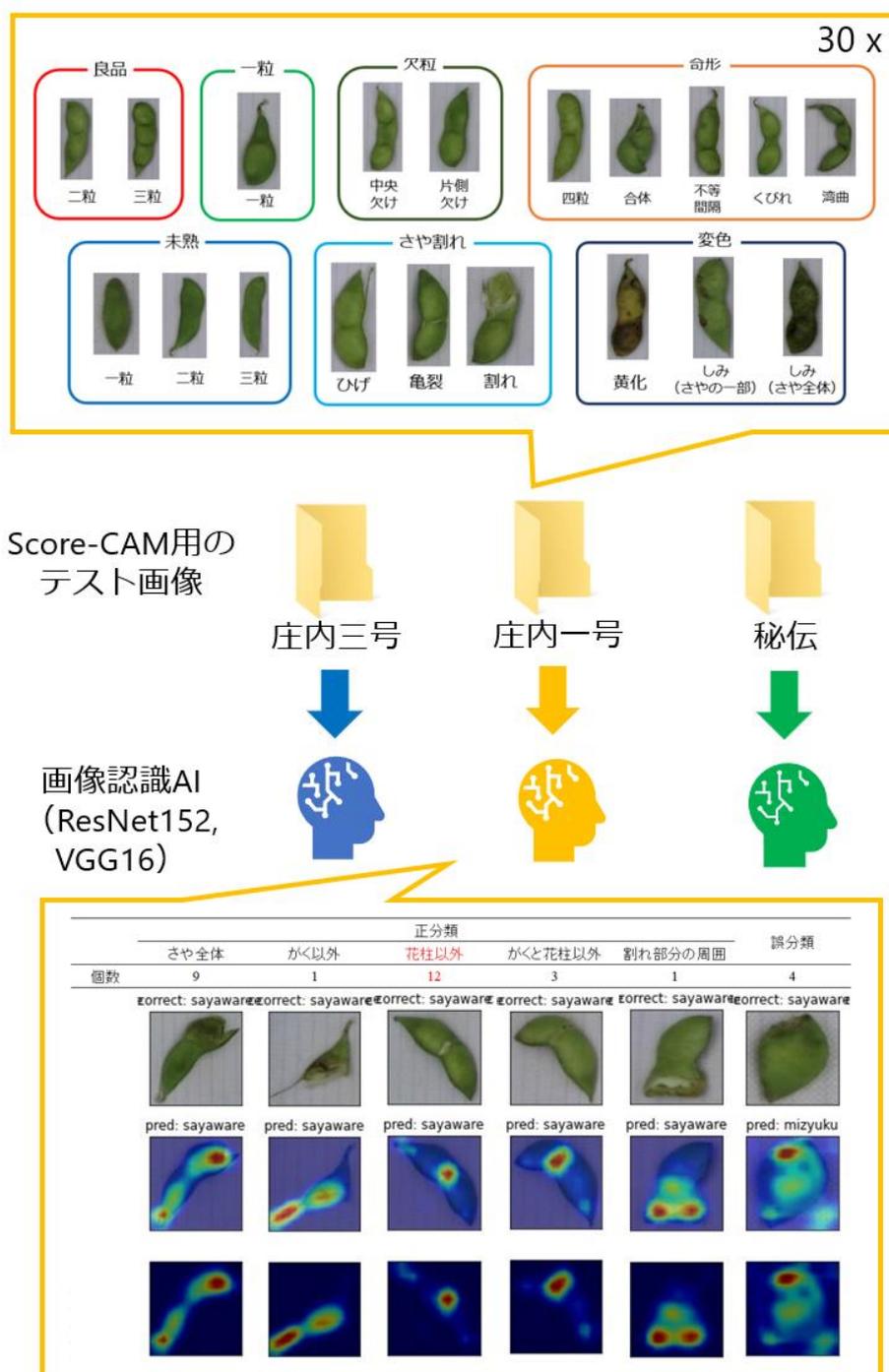


図 3-10 Score-CAM で各細分類項目の分類根拠を特定するまでの流れ

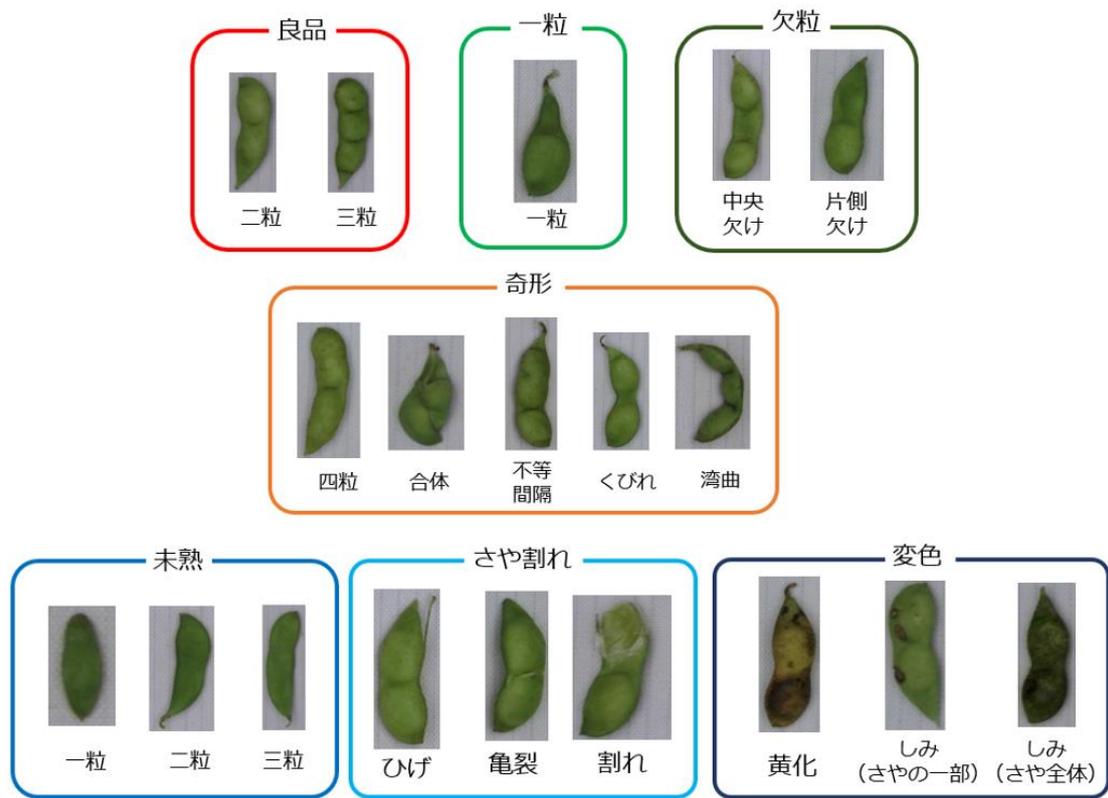


図 3-11 各細分類項目のエダマメの外観

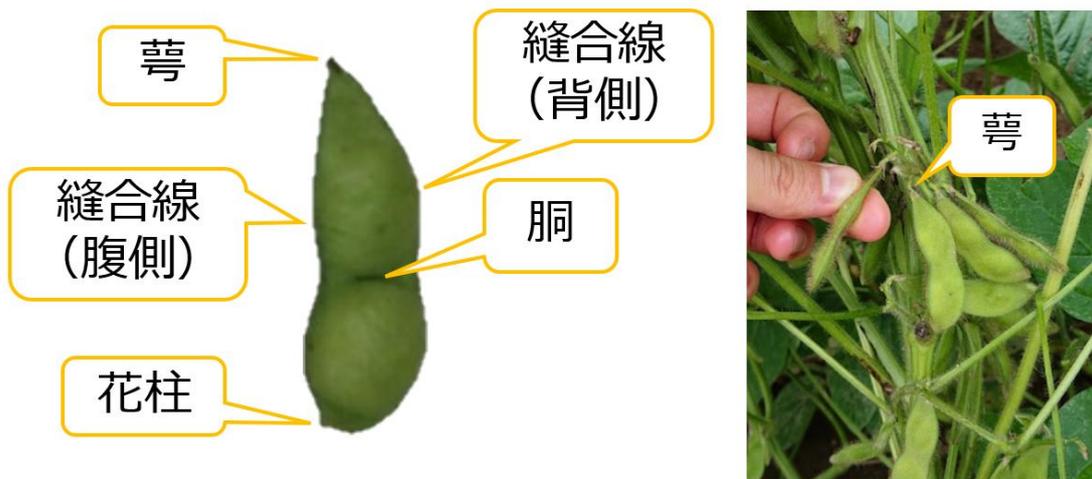


図 3-12 エダマメさやの各部位の名称

第3節 結果と考察

(1) AIの精度評価

ResNet152のAIの混同行列を図3-13~3-15に示す。各AIのAccuracyは、庄内三号と秘伝の学習データセットで0.90、庄内一号の学習データセットで0.91であった。

ResNet152の各AIの誤分類は、庄内三号が奇形を欠粒、庄内一号が奇形を変色、秘伝が変色を奇形に誤分類した数が最多であった。庄内三号で奇形を欠粒に誤分類した原因は、奇形の中に欠粒と奇形の両要素を持つエダマメが含まれていたためである。同様に、庄内一号の奇形と秘伝の変色には、奇形と変色の両要素を持つエダマメが含まれていた。また、2項目以上の要素を持つエダマメは、これらの項目以外にも含まれていた。AIは最終的に1クラスのカテゴリ結果を出力するため、2項目以上の分類要素を含むエダマメはAIで両方の特徴が抽出され、誤分類の原因となったと考える。本実験の手選別は、2項目以上の分類要素を含むエダマメを選別者の主観でいずれかの項目へ分類した。したがって、このような誤分類は第II章の結果・考察に示すとおり、選別の優先順位を決めてからAIを開発することで低減できる可能性がある。

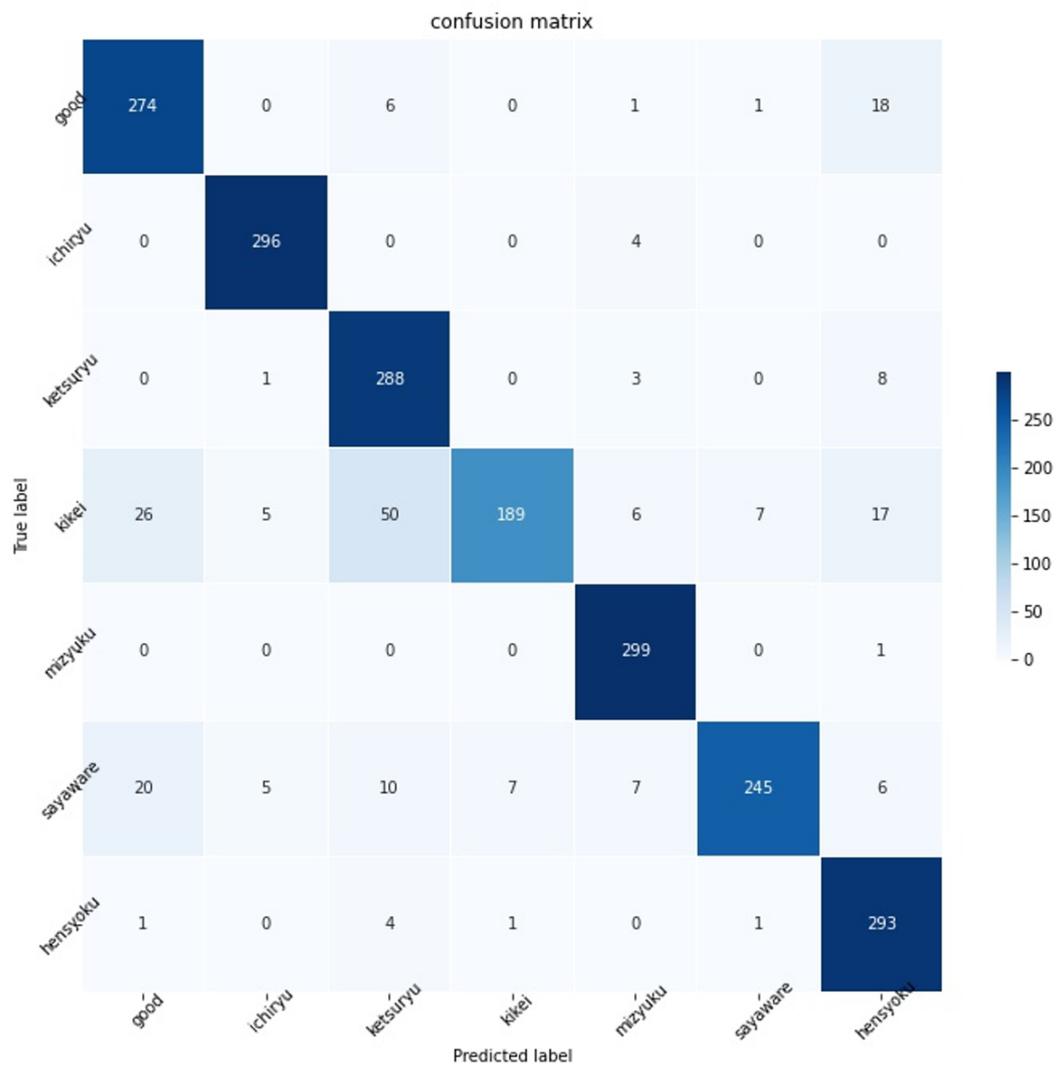


図 3-13 庄内三号の学習データセットで開発した AI (ResNet152) の混同行列

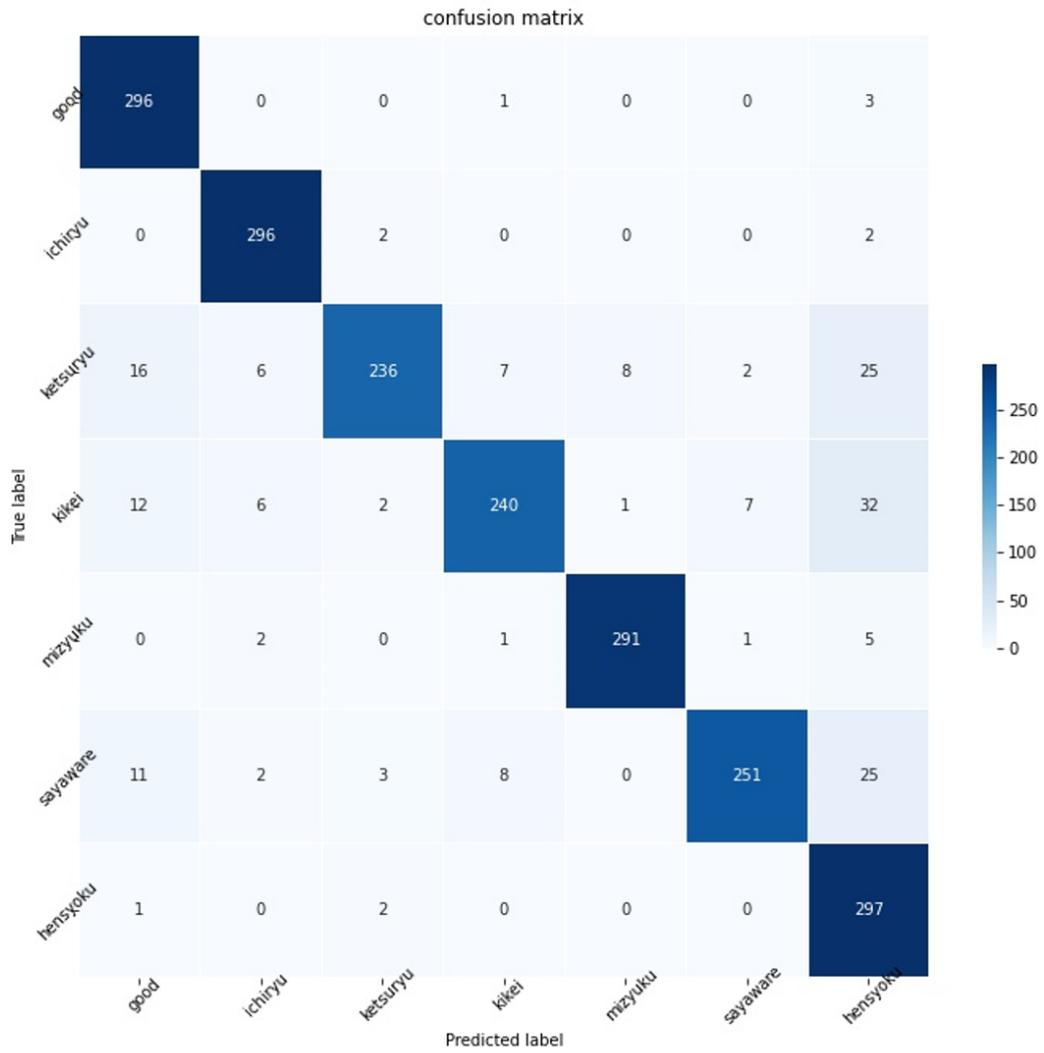


図 3-14 庄内一号の学習データセットで開発した AI (ResNet152) の混同行列

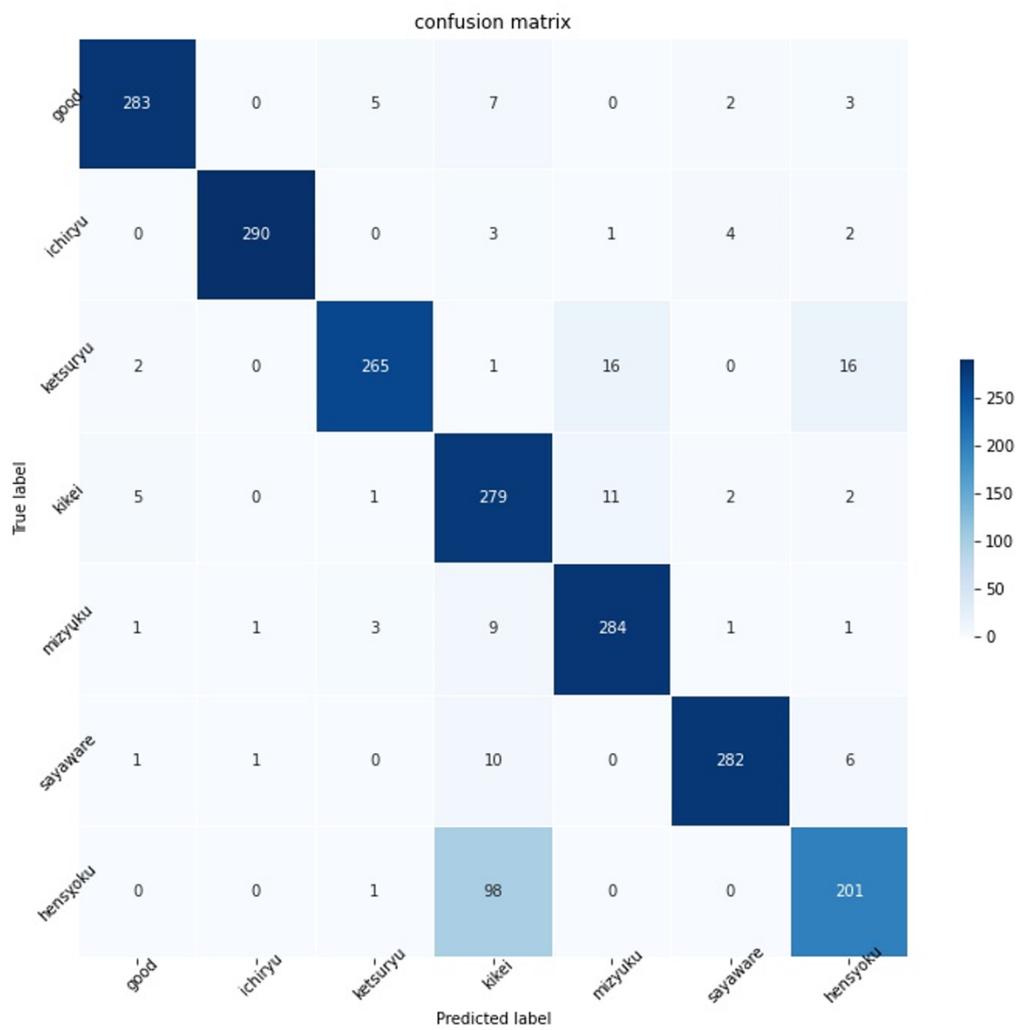


図 3-15 秘伝の学習データセットで開発した AI (ResNet152) の混同行列

次に、VGG16 の AI の混同行列を図 3-16~3-18 に示す。各 AI の Accuracy は、庄内三号の学習データセットで 0.88，庄内一号の学習データセットで 0.84，秘伝の学習データセットで 0.85 であった。

VGG16 の各 AI の誤分類は、庄内三号が奇形を良品，庄内一号が欠粒を良品，秘伝が良品を欠粒に誤分類した数が最多であった。VGG16 の AI では、良品と欠粒に関する誤分類が多かった。この原因は、第 II 章の考察と同様と考えられる。すなわち、AI は画像中のさやの写り具合から CNN で特徴抽出を行っており、さやの具体的な厚みを計測して分類していない。そのため、さやの厚さが同程度の良品と欠粒は、近似した特徴量になったと考えられる。また、ResNet152 と VGG16 の AI を比較すると、Accuracy は、いずれの学習データセットでも ResNet152 の方が高かった。ResNet152 は、VGG16 よりも CNN の層数が多いことから、各分類項目間の微細な特徴の差を学習できたといえる。

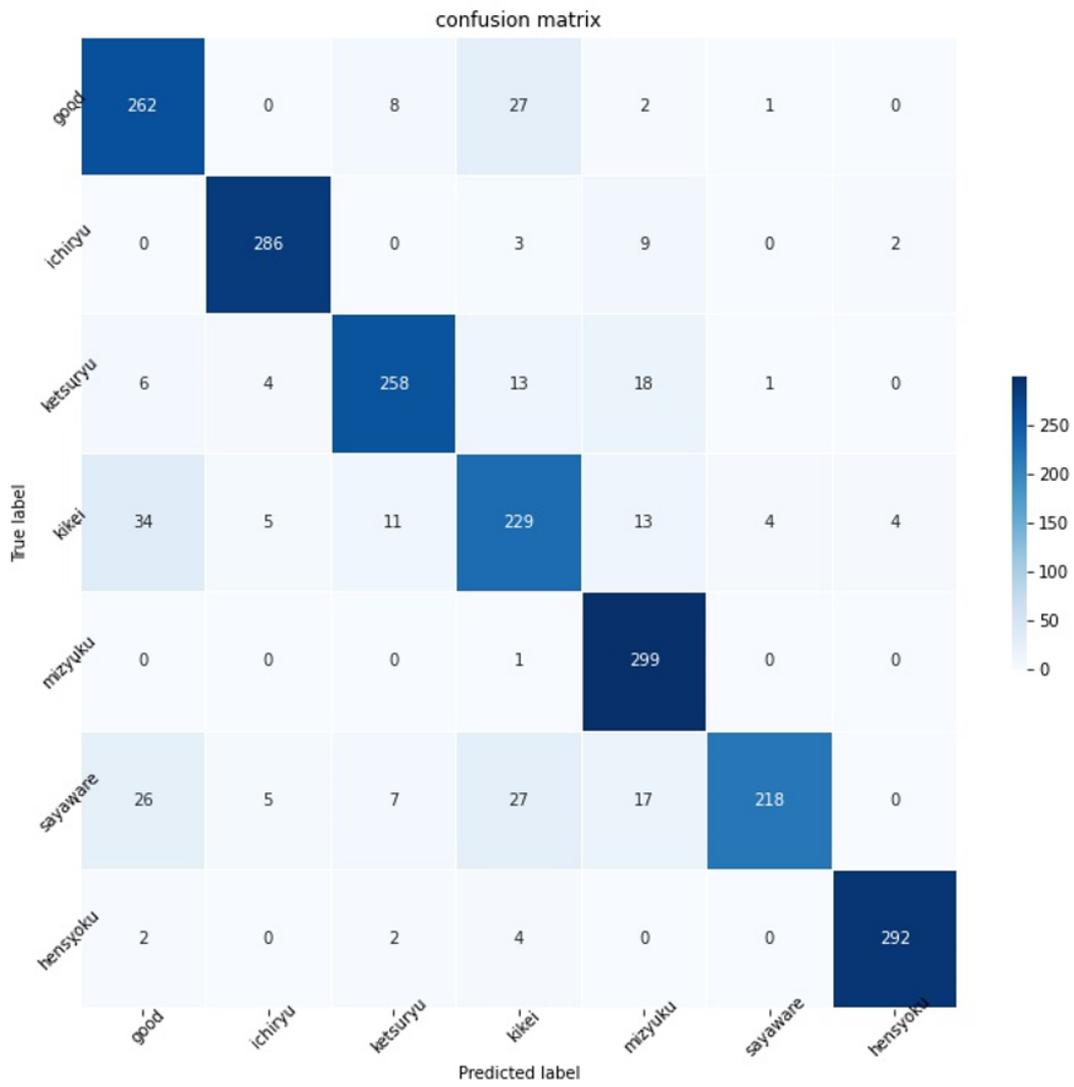


図 3-16 庄内三号の学習データセットで開発した AI (VGG16) の混同行列

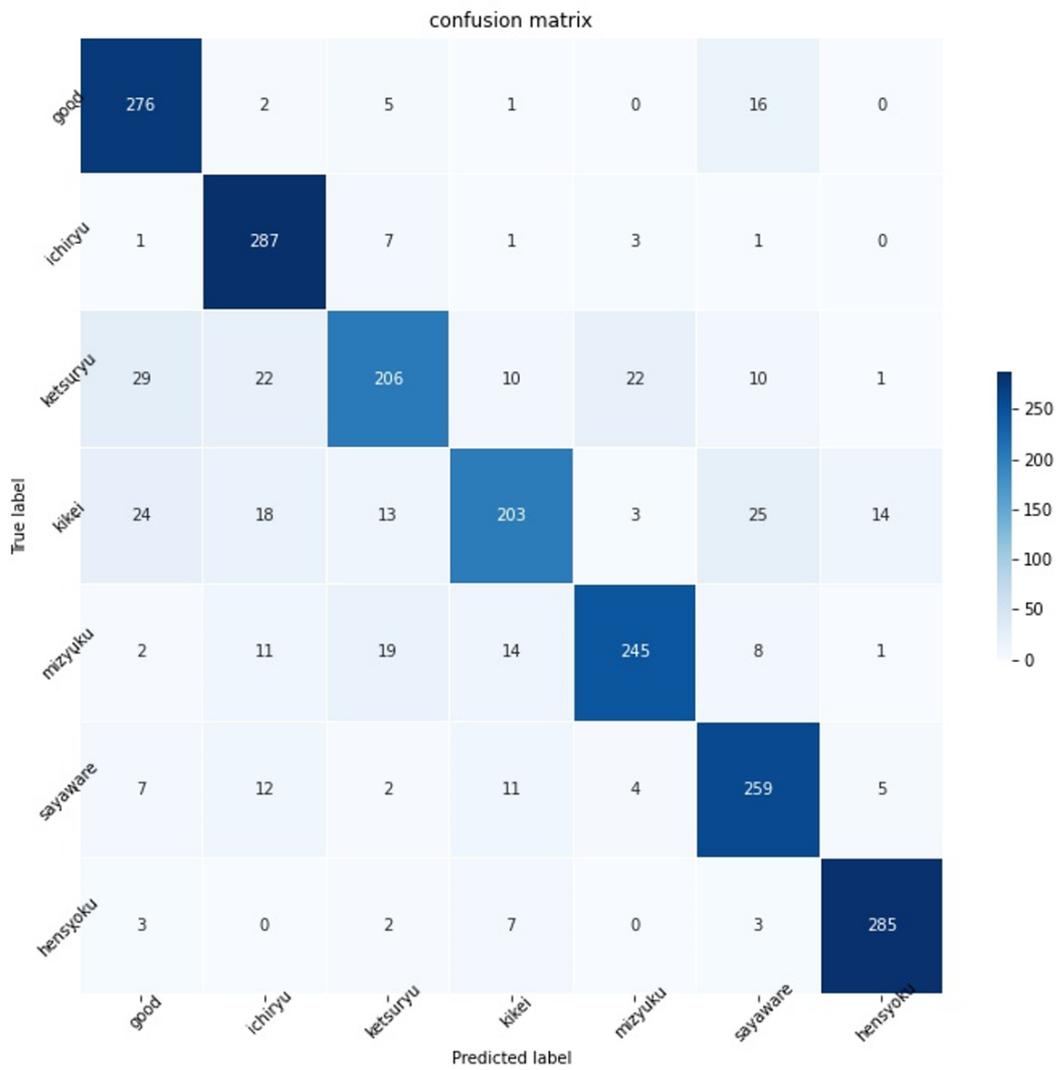


図 3-17 庄内一号の学習データセットで開発した AI (VGG16) の混同行列

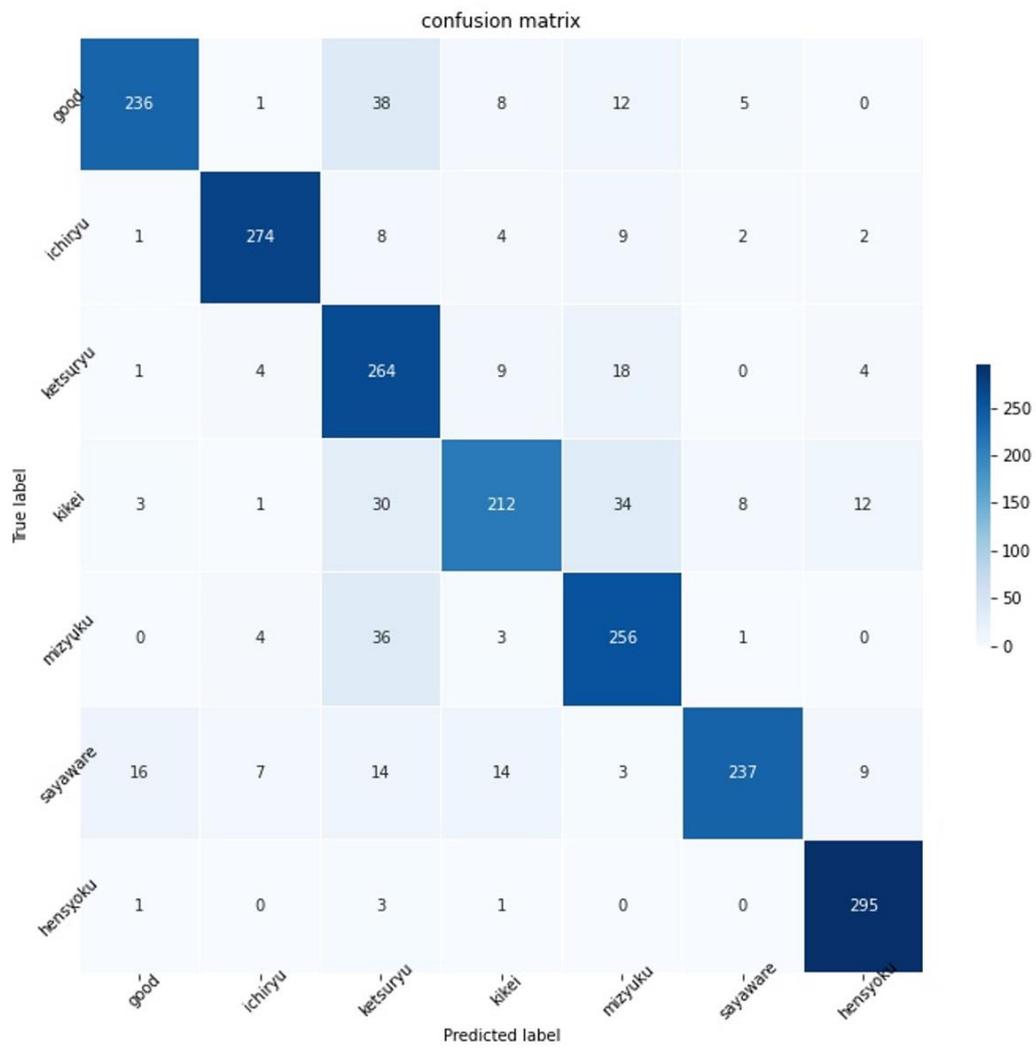


図 3-18 秘伝の学習データセットで開発した AI (VGG16) の混同行列

(2) Score-CAM による各 AI の分類根拠の特定

1) 品種が分類根拠の変化に与える影響について

Score-CAM で特定した各 AI の分類根拠を表 3-4~3-5 に示す。ResNet152 の AI に関し、品種間の分類根拠に差異があった項目は、奇形の「くびれ」と未熟の「一粒」、さや割れの「亀裂」、そして変色の「しみ(さやの一部)」の 4 項目以外であった。次に、VGG16 の AI に関し、品種間の分類根拠に差異があった項目は、良品の「二粒」と奇形の「四粒、合体、くびれ」、未熟の「一粒」、そして変色の「黄化、しみ(さやの一部)」の 7 項目以外であった。

AI が品種に関わらず、各分類項目の特徴を学習できているならば、各細分類項目の分類根拠がどの品種でも一致すると考えられる。しかし、ResNet152 と VGG16 の両方の AI は、半分以上の細分類項目で品種間の分類根拠が一致しなかった。学習データセットの相違点は品種のみであったため、エダマメの精選別用 AI が DL で各外観品質の情報と一緒に品種固有の情報も特徴量として利用しているといえる。そして、各品種固有の色合いや外形の違いは、AI にとって精度低下の因子になるといえる(以下、品種バイアス)。

また、ResNet152 の AI は、VGG16 の AI よりも品種間の分類根拠の差異が 4 項目多かった。ResNet152 は、VGG16 よりもネットワーク層数が多いことから特徴抽出能力が高く、分類精度も高いとされる⁷⁵⁾。よって、品種バイアスの影響は、特徴抽出能力の高い CNN ほど大きくなるといえる。CNN は ResNet が開発された 2015 年以降も NASNet⁷⁶⁾や EfficientNet などの新型が登場しており、CNN の特徴抽出能力が年々向上し続けている。CNN の特徴抽出能力の向上は、開発する AI の全体的な精度向上に寄与する一方で、学習データ内の精度低下因子の影響を受けやすくなると考える。よって、今後 CNN ベースの AI を開発する際には、学習データ

内の精度低下因子の特定や排除方法を確立することがより重要になる。
 それらの課題を解決することは CNN の能力を発揮させ、より高精度な AI
 開発や AI の品質保証が可能になる。

表 3-4 ResNet152 の画像認識 AI の各細分類項目の分類根拠

画像認識AIに 分類させる 外観品質項目	細分類項目	画像認識AIの開発用画像データセットに含む品種		
		庄内三号	庄内一号	秘伝
良品	二粒	花柱以外	花柱以外	さや全体
	三粒	花柱以外	さや全体	さや全体
一粒	一粒	花柱以外	花柱以外	さや全体
欠粒	中央欠け	花柱以外	花柱以外	さや全体
	片側欠け	花柱以外	花柱以外	さや全体
奇形	四粒	花柱以外	花柱以外	がくと花柱以外
	合体	花柱以外	花柱以外	合体部分とその周囲
	不等間隔	花柱以外	花柱以外	さや全体
	くびれ	花柱以外	花柱以外	花柱以外
	湾曲	さや全体	花柱以外	花柱以外
未熟	一粒	さや全体	さや全体	さや全体
	二粒	さや全体	花柱以外	さや全体
	三粒	さや全体	さや全体	花柱以外
さや割れ	ひげ	さや全体	花柱以外	花柱以外
	亀裂	花柱以外	花柱以外	花柱以外
	割れ	花柱以外	割れ部分とその周囲	割れ部分とその周囲
変色	黄化	変色部周辺	さや全体	がくと花柱を除く茶褐変部
	しみ（さやの一部）	変色部周辺	変色部周辺	変色部周辺
	しみ（さや全体）	がくと花柱以外	変色部周辺	がくと花柱を除く変色部

表 3-5 VGG16 の画像認識 AI の各細分類項目の分類根拠

AIに分類させる 外観品質項目	細分類項目	品種		
		庄内三号	庄内一号	秘伝
良品	二粒	さや全体	さや全体	さや全体
	三粒	花柱以外	さや全体	さや全体
一粒	一粒	さや全体	花柱以外	さや全体
欠粒	中央欠け	さや全体	花柱以外	さや全体
	片側欠け	花柱以外	花柱以外	さや全体
奇形	四粒	さや全体	さや全体	さや全体
	合体	さや全体	さや全体	さや全体
	不等間隔	花柱以外	がくと花柱以外	さや全体
	くびれ	さや全体	さや全体	さや全体
未熟	湾曲	さや全体	がくと花柱以外	さや全体
	一粒	さや全体	さや全体	さや全体
	二粒	がく以外	花柱以外	さや全体
さや割れ	三粒	さや全体	花柱以外	さや全体
	ひげ	花柱以外	花柱以外	さや全体
	亀裂	花柱以外	さや全体	さや全体
変色	割れ	花柱以外	がくと花柱以外	花柱以外
	黄化	さや全体	さや全体	さや全体
	しみ（さやの一部）	変色部周辺	変色部周辺	変色部周辺
	しみ（さや全体）	花柱以外	がくと花柱以外	変色部周辺

2) 品種バイアスの解消方法について

第 II 章の実験において、F 値の平均は学習データに選別対象の庄内三号の画像が含まれていない場合に有意に低下した。この現象は、品種バイアスに起因していたといえる。第 II 章の実験では、3 つのデータセットが選別対象の庄内三号を学習データに含んでいなかった。その 3 つのデータセットは B（庄内一号のみ）と C（グリーン 75 のみ）、F（庄内一号とグリーン 75）である。これらのデータセットで開発した物体検出 AI は、テスト画像の庄内三号から特徴抽出をする際に庄内一号やグリーン 75 の固有の外観情報を特徴抽出できない。これによって、それらの物体検出 AI は検出や分類に十分な特徴量を抽出できず、精度が低下したといえる。しかし、F 値の平均は YOLOv3 と Faster R-CNN の両方でデータセット B と C よりもデータセット F で有意に高くなった。データセット F の物体検出 AI の F 値は、品種バイアスの観点からデータセット B と C の物体検

出 AI と同等になるはずである。これは、複数の品種を学習データへ混合することで、AI が各分類項目の情報と品種固有の外観情報を紐づけずに特徴抽出を行うようになったためと考えられる（図 3-19）。すなわち、開発者は、複数の品種を組み合わせることで、各品種の外観的特徴を精選別に重要でない情報として AI に学習させることが可能になる。以上から、品種を組み合わせることは、品種バイアスの解消方法として有用といえる。

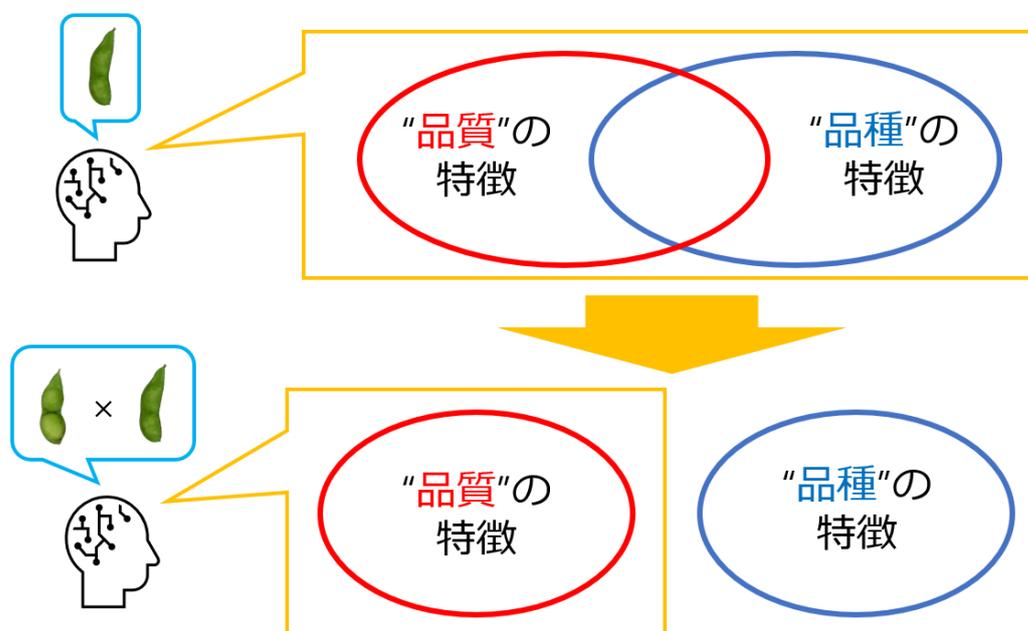


図 3-19 品種の混合による品種バイアス解消の模式図

エダマメの精選別用 AI の開発者は品種バイアスを考慮し、全品種のエダマメ画像を収集することも考えられる。しかし、エダマメは日本国内だけで 400 品種以上存在する⁷⁷⁾ため、現実的でない。第 II 章の結果から、品種バイアスは外観にわずかな差異しかない同系統の品種間でも発生した。第 II 章及び本章の実験結果から、DL で開発した AI は品種の僅かな外観の差も特徴として学習することが判明したため、1 品種の画像のみで

学習した AI がその品種の選別のみに対応する可能性が高い。また、選別可能な品種を限定することは利用者を限定するため、将来の精選別機の開発や普及を阻害する要因となる。これに対し、品種を組み合わせた AI 開発は、データ収集のコストや将来の精選別機の利用者のニーズの両方を解消する 1 つの方法になる。

3) 学習データに潜む精度低下因子の特定手法について

第 II 章では、精度評価を通じて、学習データに含む品種の違いが物体検出 AI の精度変化に影響していることを帰納的に考察した。そして、本章では Score-CAM を応用し、AI の精度低下の要因と考えられた品種の情報のみを変えた学習データセットを構築して AI を開発した。それによって、分類根拠が学習データに含む品種の違いで変化し、品種が精度低下の因子になっていることを見出した。本研究の実験手順は、DL で開発した AI の精度低下の原因を見つけ、その解決策を考える手法の 1 つになる。

今日我々は、DL でデジタルデータから自動的に特徴を抽出して高精度に検出や分類を行う AI を容易に開発できるようになった。一方で、開発者は AI に判断してもらいたい対象物に関するデータの準備やアノテーションを行う必要がある。DL で開発した AI の精度が低い場合、その原因は学習データの内容やアノテーションミスなどに起因するケースもある。Northcutt ら⁷⁸⁾は、コンピュータービジョンの分野で一般に使用される ImageNet や CIFAR-10⁷⁹⁾などの大規模データセットにアノテーションミスが多数存在することを指摘した。Northcutt らはアノテーションエラーを直した ImageNet を使用して DL で AI を開発したところ、精度がより高いとされる NASNet よりも精度が低いとされる ResNet で画像分類精度がより向上したことを確認した。この研究から、学習データの質は DL で開発した AI の精度に大きな影響を与えることを示した。DL による AI

開発プロジェクトは、一般にデジタルデータを大量に集めて学習データを構築する。そして、デジタルデータが大量に収集されるほど学習データ内のアノテーションエラーや各種精度低下因子も増加する。これにより、プロジェクトは多大な時間や労力をかけたにもかかわらず、高精度な AI を開発できずにとん挫する可能性がある。この問題を解決するため、近年の AI 開発では MLOps のように定期的な AI の精度監視や精度の改善を行う体制が整えられるようになっている⁸⁰⁾。こうした中で、開発者は早期に学習データの問題に気づき、原因の特定と解決を目指すことがより重要となる。本実験で行った帰納的な AI 精度低下因子の候補の発見から、CNN の分類根拠可視化技術によるその特定、そして特定した精度低下因子の解消の考案までの一連の流れが今後の AI の精度改善や AI の品質保証の重要な取り組みになる。

第4節 摘要

- 1) 画像認識 AI の学習データ内の品種が変化することで、CNN が分類根拠とするエダマメの部位も変化する。
- 2) エダマメの外観品質を分類する CNN 構造の AI は DL で各品質の情報と一緒に品種固有の外観情報も特徴量として利用しており、各品種固有の色合いや外形が品質分類の精度低下因子になる（品種バイアス）。
- 3) 本研究で開発した ResNet152 と VGG16 の 2 種類の画像認識 AI のうち、特徴抽出能力が高いとされる ResNet152 でより品種間の分類根拠が一致しなくなったことから、特徴抽出能力の高い CNN で精度低下因子の影響が大きくなる。
- 4) 第 II 章の実験では、物体検出 AI の精度評価に使用した庄内三号を含まないデータセットでも、複数の品種を組み合わせることで F 値が有意に向上した。このことから、画像認識 AI や物体検出 AI の学習データへ複数の品種の画像を組み合わせることは、DL で各外観品質の情報と品種固有の外観情報の特徴を紐づけずに特徴抽出を行うようになる。そして、品種バイアスの解消につながる。

第IV章 総合考察

総合考察では、エダマメ精選別用 AI の MLOps の考案や、物体検出 AI による選別アルゴリズムの考案、AI の学習データに関する研究の促進化、そして選別データの二次利用について検討する。

第1節 エダマメの精選別用物体検出 AI の MLOps の考案

(1) 本研究で得られた知見

本研究から、DL で開発したエダマメ精選別用の物体検出 AI に関して、次の結果を得た。

- ・学習データに含むエダマメ品種の違いが AI の精度変化に影響する（品種バイアス）。
- ・選別対象の品種の画像で開発した AI は、適合率や再現率、F 値が有意に向上する。
- ・品種バイアスを解消する方法は、複数の品種で学習データを構成することである。
- ・学習データに含まれる AI の精度低下因子（バイアス）の特定は、まず、AI の精度評価を通じて精度低下の因子を推測する。その後、その因子の条件のみを変えた AI に CNN の分類根拠可視化技術を使って特定できる。
- ・学習データに関する AI の精度低下の解決策は、因子の条件を変えた学習データセットで複数の AI を開発し、有意差が生じたデータセット間を比較して見つけることである。
- ・DL で開発した AI の誤分類を減らすには、優先的に選別したい分類項目を事前に決め、2 つ以上の分類要素があるエダマメに対して優先度

の高い項目をアノテーションする。

(2) エダマメ精選別用 AI の MLOps

エダマメ精選別用の物体検出 AI を開発する MLOps を以下に示す(図 4-1)。

- ①エダマメの分類項目と優先度の決定
- ②撮影方法・運用規格の統一
- ③エダマメ画像や動画の収集とアノテーション
- ④品種を組み合わせた学習データセットの構築
- ⑤DL による物体検出 AI の開発と精度評価
- ⑥運用中の物体検出 AI の精度モニタリングと精度低下時のデータ解析
- ⑦精度低下因子の絞り込みと分類根拠可視化技術による因子の特定
- ⑧因子の解消方法の検討と AI の再開発

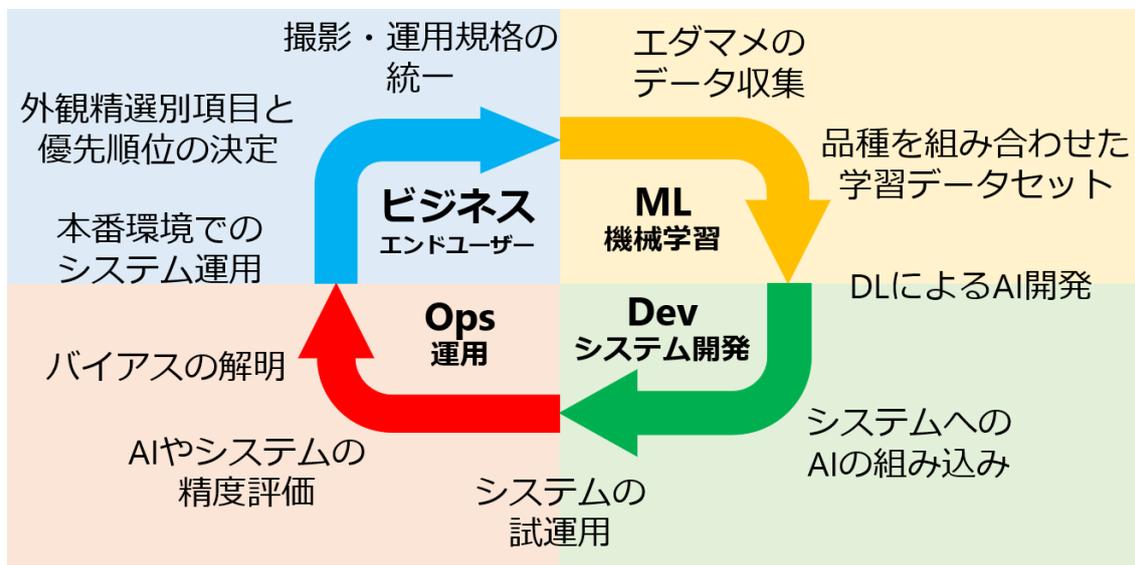
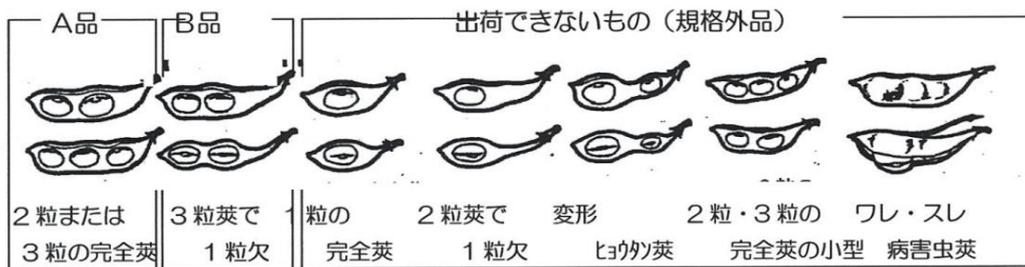


図 4-1 エダマメの精選別用物体検出 AI の MLOps の考案

①に関して、本研究で AI が誤分類した原因は、1つのエダマメに複数の分類要素が含まれる場合、ルールを決めずにアノテーションしたため

と考えられた。この対策として、分類の優先順位を事前に設定しておくことで、AIのネットワークが優先順位の高いクラスへ分類するように調整されるようになるといえる。なお、この優先順位は、A品やB品、規格外品に分類する項目を事前に検討してから決める。これは、物体検出AIの開発や運用に関係している。開発者は、AIに分類させる項目(クラス)をアノテーション時に決める。アノテーションデータは、クラスの名前や順番を統一しておく必要がある。これは、分類項目の順番や名前が一致していないと、アノテーションデータを後から共有できなくなるためである。もし、AIを開発した後に分類項目の内容や名前を変更する場合、開発者はアノテーションをやりなおす必要があり、AI開発の労力や時間を増やすことになる。開発者はこの問題を未然に防ぐため、あらかじめ各地域でエダマメ選別基準を定めてアノテーションを行う。AI開発は選別基準を統一することで学習データの二次利用が可能になり、精選別機の普及や学習データの追加がしやすくなることからAIの精度向上につながる。

分類項目の優先度について、B品や規格外品の中にも優先順位があり、それらも事前に決める必要がある。ここでは、新潟県の出荷規格⁶⁷⁾(図4-2)を例に、筆者による分類項目の優先順位付けを紹介する。選別で一番重要なことは、病原や虫の混入などで衛生的な品質を下げるエダマメを除去することである。その次に重要なことは、外観などの嗜好性を下げるエダマメを除去することである。この考えを基に、図4-3に優先順位を示す。



引用： <https://www.pref.niigata.lg.jp/uploaded/attachment/145947.pdf>

図 4-2 新潟県が決めた 3 段階のエダマメ出荷基準



図 4-3 不良品の分類優先順位

1 番目と 2 番目は衛生的な理由，3 番目以降は味や量，外観を理由に順位付けした。1 番目は，変色（虫害，しみ，黄化）である。虫害はさやの中に虫が入っている可能性があり，衛生面だけでなく異物混入として食品衛生法に違反するためである。次に，しみは変色の原因になった病原が付着している可能性があり，病原の繁殖を防ぐためである。そして，黄化は外観の悪さだけでなく，ダイズになりかけているため味が劣化しているためである。2 番目はさや割れである。これは，腐生菌が割れた部分からさやの中に入り，可食部を直接汚染する可能性があるためである。また，さやの中は閉鎖的で湿度が高く腐生菌の繁殖源になり，ほかのエダマ

メへの伝染を防ぐためである。3番目は未熟である。未熟は良品と比べて味や食べ応えが劣り、消費者の嗜好に沿わないためである。4番目は一粒である。一粒の子実1個は良品と比べて同じ重量でも数が少なく、消費者にとって損をしたと考えられるためである。5番目は奇形である。奇形は外観が良品とかけ離れており、食べづらさや不快に感じる可能性があるためである。しかし、奇形は2粒以上入っており、子実の量は良品と同等になることから一粒よりも優先順位を下げた。6番目は欠粒である。欠粒は未熟子実が混ざっているが、形状は良品と似ており、良品子実が2粒入っている個体もある。欠粒は、規格外品より等級の高いB品にも該当するため、優先順位を最も低くした。

②に関して、物体検出AIの精度を高めるには、データ収集環境と運用環境を統一する。DLで開発したAIや従来の画像処理では、照明や背景の違いがAIの精度へ影響することが報告されている⁸¹⁾ため、撮影環境の違いもAIの精度変化の因子になる。背景の変化によるAIの精度変化に関して、Ferentinosら⁵⁴⁾は、PlantVillageデータセットを利用して、植物の葉の病気を分類する画像認識AIを開発した。この実験で、Ferentinosらは、屋外で撮影した画像を使って画像認識AIを開発し、屋内で撮影した画像をテスト画像に使用した。その結果、画像認識AIの分類精度は約68%になった。一方、Ferentinosらは、屋内で撮影した画像を使って画像認識AIを開発し、屋外で撮影した画像をテスト画像に使用した。その結果、分類精度が約33%となった。このことから、DLで開発したAIは、使用環境と異なる背景の画像で開発すると、特徴量の抽出内容が変化して精度が低下する。しかし、この研究で開発された画像認識AIは、屋外と屋内の両方で撮影された画像を使うことで、分類精度が約99%になった。このことから、様々な背景パターンを学習データに含むことで、様々な背景でも高精度に分類できるAIを開発できるといえる。これは、本研

究で複数の品種を混合することで、物体検出 AI の精度を向上させたことと同様の事例である。しかし、背景の種類が多岐にわたると、収集データの背景が特定の種類に偏ることがある。この偏りが将来の AI の精度低下因子になる可能性があるため、背景を統一することが望ましい。エダマメ精選別機は、屋外で使用する農作物の収穫機や病害の画像診断 AI とは異なり、人為的に背景を統一できる。そのため開発者は、学習データと AI の利用環境の照明や背景を統一し、エダマメの外観以外の変化要因を極力なくすとよい。

また、照明やカメラの設定も統一する必要がある。照度や照明の配置は、AI がエダマメの未熟度を画像から判別できるように、子実部周辺に影ができるように設定するとよい。次に、カメラは、どのカメラでもエダマメを同じような彩度やコントラストで撮影できるように共通のキャリブレーションを行う。

③と④に関して、本研究より、品種が物体検出 AI の精度に影響する因子であることが分かった。よって、あらかじめ品種を混合した学習データセットを作成し、⑤で物体検出 AI を開発して運用する。

⑥に関し、AI の精度に影響する因子は、まだ存在すると考えられる。例として、栽培条件によって変化する同一品種の外観の違いである。農作物は同一品種であっても栽培地域や栽培時期、栽培方式などの変化により外観が変化する。本研究の実験で使用したエダマメは山形県鶴岡市の畑地ほ場で栽培したものであり、第 2 章と第 3 章の実験はそれぞれ 2019 年と 2020 年のエダマメのみを使用した。そのため、本研究では栽培地域や栽培時期によるエダマメの外観変化が物体検出 AI の精度変化に及ぼす影響を調査していない。今後、各栽培条件の相違による同一品種の外観変化と分類精度の変化に注目する。

⑦と⑧に関し、⑥で新たな因子による物体検出 AI の精度変化が確認さ

れた場合、第 3 章の実験手法に基づいて、精度低下の因子と考えられる要素のみを変えた複数のデータセットで画像認識 AI を開発する。そして、Score-CAM などの分類根拠可視化技術による分類根拠の変化を確認して、因子の特定とその解消方法を考案すべきである。

以上の点に注意して AI 開発やその運用を行うことで、高精度なエダマメ精選別機を実現し、最終的に精選別作業の能率向上や人件費の削減、処理量の増加による出荷額の増加で生産者の収益向上につながる。物体検出 AI は、年々新モデルの登場や精度の改善が進んでいる。本研究で使用した YOLOv3 の登場以降、改良モデルである Scaled-YOLOv4⁸²⁾や YOLOR⁸³⁾、YOLOv7⁸⁴⁾などが登場し、MS COCO データセットによる精度比較では、従来の物体検出 AI よりも精度が向上したことが報告されている。AI の特徴抽出能力や検出・分類能力が向上することで、AI が学習データに含む精度低下因子の影響を受けやすくなり、それに伴い学習データの内容がより重要になる。これは、エダマメの精選別に限らない。各用途の AI を DL で開発する場合、学習データのどのような内容が精度低下因子になるか検討し、その対策とデータ整備にも気を付ける必要がある。

第 2 節 DL で開発した物体検出 AI とオープンソースソフトウェアを搭載した精選別システム

本研究で使用したフレームワーク「Darknet」は、C 言語で実装されたオープンソースソフトウェアである。Darknet は「YOLO LICENSE Version2」というオープンソースライセンスでリリースされており、無料で利用可能であるため、自分の用途に合った物体検出 AI を容易に作成できる。すなわち、AI の開発者は、本研究の考察を基にエダマメの画像を収集し、物体検出 AI に分類させたい項目を設定してアノテーションを行えば、自分の目的に合った物体検出 AI を開発できる。従来、このような画像処理ソフトウェア（AI）は高度なプログラムの知識を有する専門家のみ作成可能であったが、今では画像を収集してアノテーションを行えば、Darknet が DL を行い自動で物体検出 AI を開発する。Darknet による DL の実行や開発した物体検出 AI の稼働は、短いコマンドを PC へ入力するだけで容易に行える。さらに、Darknet はカメラで撮影した動画をリアルタイムで物体検出 AI へ入力し、フレーム単位で物体検出 AI が検出・分類した内容を JSON 形式でリアルタイム出力する機能がある（図 4-4）。したがって、Darknet とそれから出力されたエダマメの座標情報を基に不良品除去部分へ信号を送るプログラムを組み合わせれば、物体検出 AI を搭載した高精度なエダマメ精選別機を開発できる。本節では、Darknet を中核としたエダマメ精選別機のシステムを考案する。

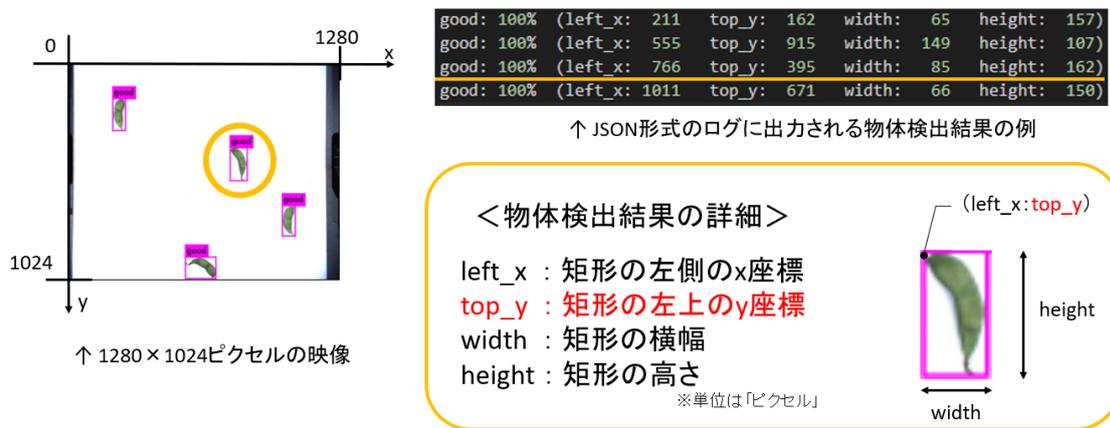


図 4-4 物体検出 AI が処理した検出・分類結果を Darknet がログデータとして JSON ファイルへ出力した内容

Darknet を利用したエダマメ精選別機のシステム概略を図 4-5 に示す。精選別機はベルトコンベヤによるエダマメ搬送部と検出部，Darknet と物体検出 AI などを含む情報処理部，そして不良品除去部で構成する。

最初に，エダマメは検出部のカメラで撮影され，その動画が Darknet を介して物体検出 AI に入力される。次に，物体検出 AI は動画をフレーム単位で処理し，フレームごとのエダマメの座標と分類結果を出力する。その出力情報は Darknet が JSON 形式のログデータとして出力する。その JSON 形式のデータを Darknet から受け取り，その中から特定の y 座標の範囲内にあったエダマメの情報を抜き出すプログラムが Darknet と連動して稼働する。物体検出 AI は動画をフレーム単位で物体検出するが，前フレームの物体検出結果とはリンクしていないため，搬送部を流れていくエダマメを同一物として追跡できない。そのため，提案方式では特定の座標に到達したエダマメの情報を使用する。この特定の y 座標内に不良品へ分類されたエダマメがあった場合，そのエダマメの x 座標情報を基に，その x 座標をカバーしている不良品除去部へ信号を送る（図 4-6）。エダマメの精選別は，この一連の流れで行うことができる。また，特定の

y 座標内で検出された各分類項目のエダマメをカウントして記録することで、選別データを選別以外の用途に二次利用できる。

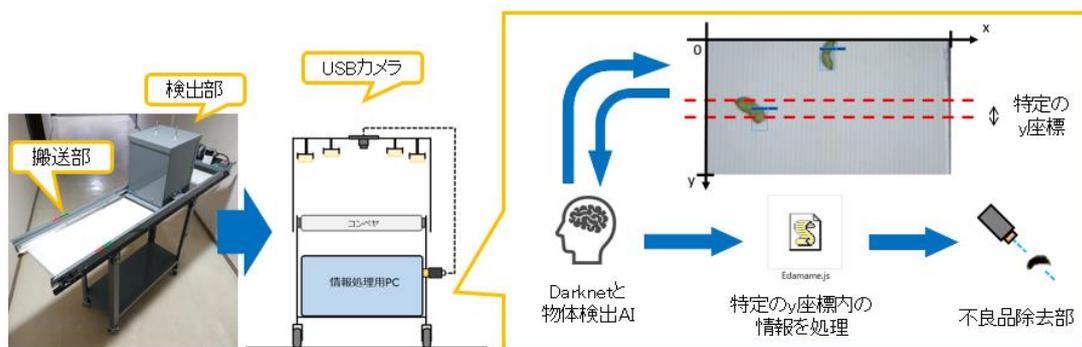


図 4-5 Darknet と物体検出 AI を利用したエダマメ精選別機のシステム概略

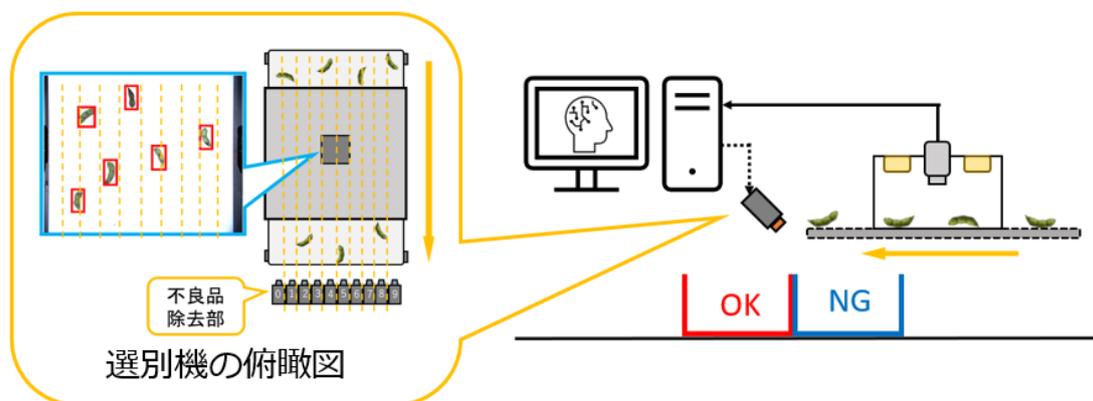
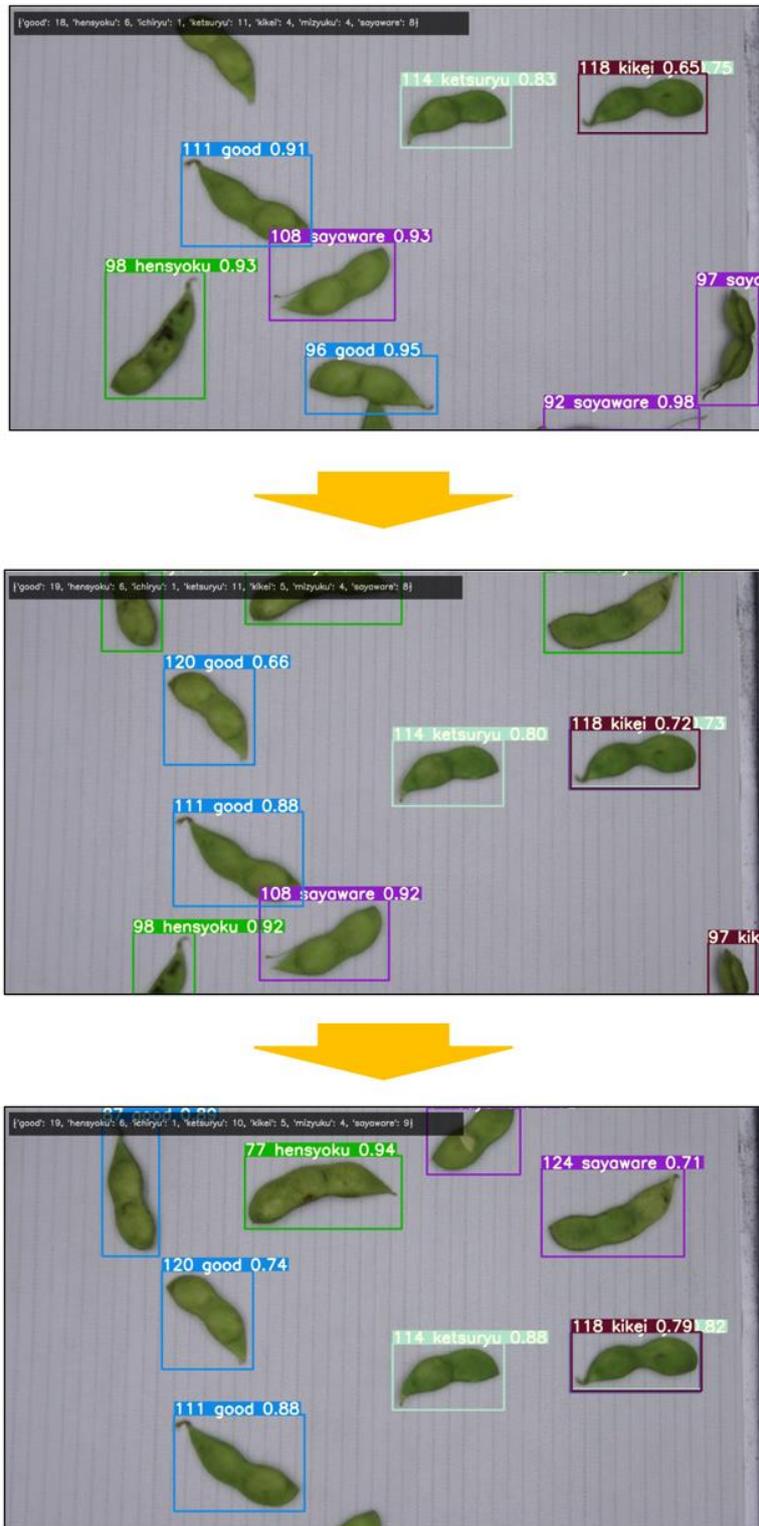


図 4-6 不良品除去部周辺のシステム構成

提案方式では、物体検出 AI が搬送部を流れるエダマメを同一物として認識できないことから、特定の y 座標内で検出されたエダマメの情報を利用する。この方式の課題は、特定の y 座標内でエダマメを未検出することや重複検出する可能性があることである。そのため、映像の端から端

までを流れるエダマメを同一個体として追跡することができれば、検出漏れや重複検出を低減できる。近年、物体検出 AI と前後フレームの物体が同一物か判断する再同定 AI (ReID モデル) を協働させ、動画内を移動する対象物を同一物として追跡可能な MOT (Multiple-object tracking) アルゴリズムが考案された。主な MOT アルゴリズムには、DeepSORT⁸⁵⁾がある。DeepSORT は最初に YOLO などの物体検出 AI で対象物を検出し、次のフレームで検出された検出物の位置情報と ReID による前後フレームの検出物の同一性を評価して追跡を行う (図 4-7)。

MoT による搬送部上を流れるエダマメの個体追跡やカウントは、先述の提案方式よりもコンピューターの GPU メモリの使用量が増加する。2022 年現在の GPU で MoT をリアルタイムで処理するには、演算コストの高さから高額な GPU を必要とする。そのため、選別システムには、マシンスペックの低さやリアルタイム性を重視する場合に MoT よりも負荷の少ない提案方式を採用すべきである。しかし、GPU の性能やアルゴリズムの進化は日々進んでおり、今後はより安価で高性能なコンピューター上で MOT をリアルタイムで稼働させることが容易になる。よって、選別アルゴリズムには現在のマシンスペックを基準として考えずに、MOT の利用も検討すべきである。



ソースコード：Mikel Broström, Real-time multi-camera multi-object tracker using YOLOv5 and StrongSORT with OSNet⁸⁶⁾

図 4-7 MOT アルゴリズムによる搬送部上を流れるエダマメの物体追跡と各分類項目のカウンタ例

第3節 学習データの内容・構成に関する研究の促進

学習データの内容に関する研究は、エダマメの精選別に関わらず、まだ十分に行われていない。これは、第1章で説明したタイプ2の研究者の少なさに起因していると考えられる。また、タイプ2の研究者は、用途に応じた学習データを実験ごとに収集していることも原因といえる。それに対して、タイプ1の研究者が多い理由は、ImageNetなどの共通で利用できる学習データセットが整備されており、さらに誰でも自由に利用できるためと考えられる。このように、今後は現場で収集されたデータが公開され、誰でも利用できる仕組みを作っていくことも大切である。

この仕組みの利用は、タイプ2の研究が活性化すると予想され、新規のAI研究者が独自にデータを収集しなくてもAIを開発できる。もしくは、公開されている学習データと新規の学習データを比較または組み合わせ、学習データの問題点やその改善が行われやすくなる。

また、学習データの公開は、研究の信頼性向上や次の研究に役立つといえる。タイプ1の研究が活発である別の理由に、開発されたAIモデルがインターネットで公開されていることもあげられる。開発されたAIを自由に利用できることは、第三者がそのAIの性能を体験でき、それをベースに次の研究に繋げやすくなるためである。例として、本研究で使用したフレームワーク「Darknet」とYOLOが挙げられる。YOLOの開発者であるRedmonらは、開発したDarknetとYOLOを彼らのホームページ⁸⁷⁾に公開した。その後、様々な研究者が、YOLOをベンチマークにして、新型の物体検出AIを開発した。そして、BochkovskiyらがRedmonらの後を継ぎ、Darknetの改善や新型のYOLOを開発し続けている。筆者は、第三次AIブームが盛り上がっている理由は、こうしたデータの積極的な公開と二次利用が活発であるためと考えている。これまでは、各研究の成果物が非公開であることがほとんどであった。現在はデジタルデータを積極

的に公開して、研究成果をより広めて発展させる取り組みが活発になっているため、学習データも既存の考え方にとらわれずに積極的に公開して共同で改善をすることが求められる。

第4節 選別情報の二次利用

物体検出 AI を搭載したエダマメ精選別機は、搬送部を流れたエダマメの数やその分類項目の選別情報を記録できる。この情報を用いて、精選別だけでなく、他の業務や栽培の改善につなげていく方法を検討する。筆者が考案した選別情報の二次利用は、次の通りである。

- (1) 生産者と運送業者の各ニーズを満たす集荷システム
- (2) 日ごとの収穫量の調整
- (3) 選別項目の自動設定
- (4) 栽培へのフィードバック
- (5) 育種への応用
- (6) 精選別以外の AI との連携

(1) 生産者と運送業者の各ニーズを満たす集荷システム

生産状況によって変わる可能性があるが、生産者は品質を優先するタイプとコストを優先するタイプの 2 つのタイプが存在する。エダマメは収穫後に常温で保管すると、風味や色合いの劣化が早まる。品質が低下したエダマメは、市場価格が低下する。そのため、エダマメの品質をより長く保ちたい生産者は、出荷調製後にできるだけ早く運送業者へ出荷しようとする。一方、輸送費を削減して経費を抑えることを優先する生産者は、集落でまとめて集荷を依頼することや、運送業者の都合に合わせて出荷して輸送費を圧縮しようとする。いずれの生産者も運送業者と連絡を取り、集荷のタイミングや輸送費を調整している。

運送業者は直接生産者の元へ出向くことや電話などで連絡をとり、その日の集荷時間や量を確認している⁸⁸⁾。運送業界では、農業分野と同様にドライバーの人手不足が深刻化している⁸⁹⁾ため、配車や運搬量の最適化が運送業者にとって輸送費や車両費、人件費などのコスト削減、温室効

果ガスの発生を抑えることにつながる。

選別データのリアルタイム集計は、目標出荷量の把握と選別終了時間を予測し、それぞれのニーズを加味した最適な集荷ルートを自動作成できるようになる可能性がある（図 4-8）。データセンターは各精選別機から選別の終了予測時間を受け取り、生産者の嗜好に応じた重みづけを作成する。この場合、輸送コストも併せて可変できる仕組みも必要である。この仕組みは、品質を落とさないスピーディな集荷とコストパフォーマンスの高い効率的な集荷につながり、2タイプの生産者のニーズに応えられる可能性がある。運送業者は手配するトラックの台数や集荷タイミングの最適化が実現することで、ドライバー不足の解消や配送時間の短縮による温室効果ガスの削減につながる。

なお、この実現には各生産者が選別状況のデータを送出可能な精選別機を導入し、送出されたデータを解析して集荷ルートを最適化する仕組みがあることを前提としている。一部の生産者の精選別機から選別データを活用するだけでは、流通全体の改善に寄与しない。そのため、集荷の最適化をするには、より多くの生産者の選別状況をリアルタイムに集約できる仕組みを構築する必要がある。

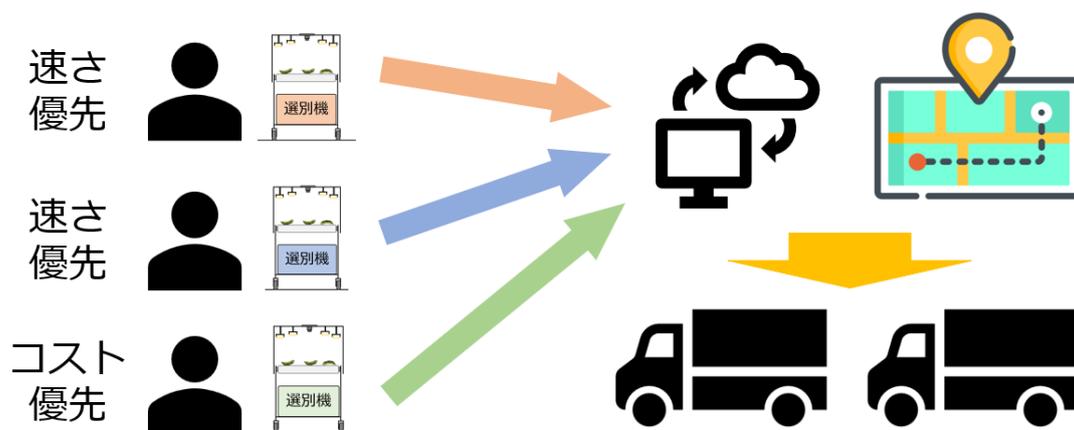


図 4-8 生産者と運送業者の各ニーズを満たす集荷システムのイメージ

(2) 日ごとの収穫量の調整

生産者は、集荷までの限られた時間で効率的に出荷調製を行っている。生産者は、収穫後の出荷調製で予定出荷量に満たない場合に追加で収穫することがある。追加で収穫する場合は、出荷に間に合わせるために早めにその判断をする必要がある。精選別機で良品の選別量をリアルタイム把握することは、出荷調製量の過不足を早期に判断することが可能となり、追加収穫の迅速な判断に寄与できる。早期の判断は他作業との兼ね合いや作業者の割り当てを効率化するため、出荷調製作業全体の改善につながる。

(3) 選別項目の自動設定

生産者は、多少の商品の品質を落としても良品の基準を変えて選別することがある。これは、必要な収益の確保や、市場の要求に合わせて出荷量を調整するためである。

ある程度の利益を確保するための出荷量は、その時の市場価格や需要量を把握することで推測できる。しかし、生産者が常にそうしたデータを手に入れるとは限らず、どの時期にどの程度の量を出荷したらよいかを把握できれば選別基準を考えやすくなる。これらの情報を得るには、流通業者の協力が必要になる。流通業者は、消費者の需要に合った量のエダマメを安定的に仕入れたいと考えている。これは、売れ残りや品不足による販売ロスを減らして、安定的な収益を確保したいからである。そのため、理想は、流通業者が需要量（仕入れたいエダマメの量）を設定し生産者へ伝え、生産者がそれに応じて選別基準を決めて出荷量を調整することである。

精選別機で選別基準を緩和する場合、生産者が目標出荷量に達するように自分で選別基準を設定する必要がある。しかし、DL で開発した物体

検出 AI は、分類項目間の分類パラメータを自由に変えられない欠点がある。そのため、DL で開発した物体検出 AI を搭載した精選別機は、選別内容を緩和したい場合、不良品として除去する分類項目を変更することで対処できる。それに対して、従来の画像処理技術は、特徴量の抽出パラメータや分類閾値を変色や形状に合わせて手動で変えることで、選別基準を自由に調整できるメリットがある。2022 年現在に販売されているエダマメ精選別機の中には、利用者が手動で分類閾値を調整できる機種がある⁹⁰⁾が、生産者が画像処理の分類閾値の変更を行うのは容易でない。また、収穫したエダマメの良品・不良品の割合はほ場内で一定でない。そのため、生産者は物体検出 AI や画像処理でも定期的に選別項目や閾値を変更する必要がある。これは、生産者が常に精選別機を監視する必要があり、出荷調製作業の効率を下げることになる。よって、物体検出 AI を搭載した精選別機には、選別結果から出荷量を予測して、採算のとれる良品の量に到達しない場合に動的に選別項目を変更可能な機能が必要である。

これに対して、図 4-9 の仕組みが考えられる。この仕組みは、データセンターが全国の精選別機から収穫量や良品・不良品の割合の推定値を収集し、エダマメの合計出荷量を予測する。そして、小売市場の需要量と出荷予測量の差を計算し、出荷要求量のデータを各精選別機へ送る。各精選別機は、このデータを基にその日の収穫量や選別内容を解析し、選別項目を変える。これにより、生産者は選別状況を把握して分類項目を変更する手間がなくなり、市場の需要に合った出荷を行える。

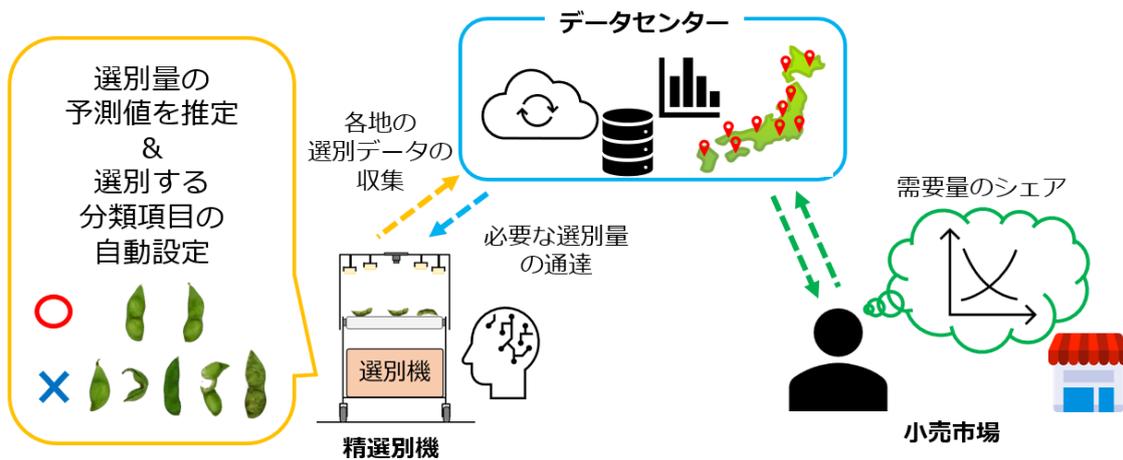


図 4-9 物体検出 AI に選別させる分類項目の自動設定システム

この提案システムは、SDGs（Sustainable Development Goals：持続可能な開発目標）⁹¹⁾にも貢献できると考える（図 4-10）。SDGs は 2030 年までに持続可能でよりよい世界を目指す国際目標であり、17 の目標と 169 のターゲットから構成される。17 の目標のうち、提案システムは次の 5 つの目標にあてはまる。

- ・ 2. 飢餓をゼロに
- ・ 7. エネルギーをみんなに そしてクリーンに
- ・ 9. 産業と技術革新の基盤をつくろう
- ・ 15. 陸の豊かさを守ろう
- ・ 17. パートナリーシップで目標を達成しよう

この提案システムは、生産者だけでなく、輸送業者や流通業者も恩恵を得られる。流通業者は、エダマメが豊作の場合に市場が必要としないエダマメを輸送する必要がなくなるため、燃料を削減できる。これは、SDGs の 7 番を達成できる。小売業者は、需要に合った量のエダマメを安定的に確保できるようになるため、売れ残りや品不足による販売ロスを減らせる。これにより、生産者への支払い金額を今までより正確に予測できるようになり、生産者の収益の安定化や栽培計画の立案にも役立つといえ

る。そして、生産者が使用する肥料や資材の節約に繋がり、SDGsの15番を達成できる。今後生産者が必要以上の肥料や資材を使わなくなれば、肥料や資材が不足している国や地域へ分配され、各地で食料生産量が増えるようになる。これにより、SDGsの2番を達成できる。最後に、提案システムは、生産者と小売業者の協力が重要である。よって、提案システムが有効に機能すれば、異業種の連携と新しい流通の仕組みが出来上がり、SDGsの9番と17番を達成できる。市場の状況を加味して選別基準を自動で決める取り組みは、将来このように広い範囲に貢献できるようになる。



【引用】 https://www.unic.or.jp/activities/economic_social_development/sustainable_development/2030agenda/sdgs_logo/

図 4-10 SDGs の 17 の目標一覧

なお、生産者の中には、目標出荷量をあらかじめ決めている人もいる。そのような生産者にとって、選別基準の自動設定は本提案システムの一部のみで十分といえる。このタイプの生産者は、事前にその日の目標出荷量を精選別機に設定する。そして、選別データの内容から予測出荷量を自

動で計算し、目標量に達しない場合は選別で不良品と判別されたものの一部を良品として出荷するようにする。このように各生産者のニーズによりシステムを使い分けることで、選別基準の自動決定と目標量の出荷ができるようになる。

なお、提案システムの課題としては、小売業者の需要量の決めづらさがある。小売業者の需要量は、消費者の需要を把握して決める。消費者の需要は時期や販売圏の人口など様々な要因が絡むため、厳密な需要予測を立てることが難しい。しかし、今後 AI は需要予測の分野にも使用され、予測精度が改善され、提案システムが実現すると考える。

(4) 栽培へのフィードバック

精選別機では場毎の収穫量や良品・不良品の割合を算出することで、ほ場ごとの栽培品質を数値化できる。これは、栽培の成績表を作成することと同様である。これまでの農業は、栽培品質に関するデータ収集や他人と同じ基準で栽培品質を比較することが少なかったため、自分の栽培の良し悪しや目標設定がしづらかった。もし、自分とほかの生産者の栽培品質を比べられれば、自分の栽培品質の良し悪しが分かる。そして具体的な生産の数値目標ができる。このように、選別データを活用することで、生産者が他の生産者と比較して栽培の問題点を整理し、次年度に向けて栽培計画や栽培改善をできるようになる。

また、精選別機で収集したデータを土壌水分計や気温計などの各種環境センサで収集したデータや気象情報などの他のデータと紐づけることで、施肥計画の改善ができる可能性がある。これを実現するには、複数のセンサや選別機から得られたデータを統合して総合的に分析を行う仕組みの構築も必要である。

(5) 育種への応用

農作物は、味や外観などの品質や生産性向上のための品種改良が盛んに行われている。調査員は、育種した品種ごとの収穫量や良品・不良品の構成割合のデータを算出するために、手選別する必要がある。手選別は育種した品種や反復が増えるほど手間が増えるため、時間を多く浪費する。育種の研究所でも精選別機を導入することで、収穫量や良品・不良品の構成割合の算出を自動化できる。そして、良品出荷率の高い系統、3粒さやの多い系統、大きさや系統の抽出といった育種研究の推進に繋がる。

(6) 精選別以外の AI との連携

精選別機から得られる選別情報は、ほ場レベルでの収穫量や病害虫に関する不良品の発生割合を数値で評価できる。農業関連の AI では、収穫量予測 AI、収穫物の病害虫の被害度推定 AI、土壌病害診断 AI などがある。これらの AI は最終出力が収穫量や良品・不良品の割合であり、選別情報との関連性が高い。よって、選別情報のデータは、収量予測 AI や病気の被害度推定 AI の正解データとして利用できる(図 4-11)。すなわち、AI が AI を進化させ、農業分野に関連する AI の開発が促進されるようになると思う。近年、栽培期間中にマルチコプターや人工衛星などでリモートセンシング⁹²⁾を行い、収穫量や農作物の病害被害度などを AI で推定する研究が活発である⁹³⁾⁹⁴⁾。こうした AI の開発では、調査員が正解データとして実際の収穫量などを調査する。人力で収穫量などのデータ収集をするには労力と時間を要するため、正解データを十分に収集できない可能性がある。この課題に対して、精選別機から得られる収穫量のデータを正解データとして利用することで、少人数で迅速に収穫量のデータを収集できる。そして、収穫量に関する各 AI の開発が促進され、さらに入力データと正解データの量が増えることで AI の精度向上も実現できる

可能性がある。

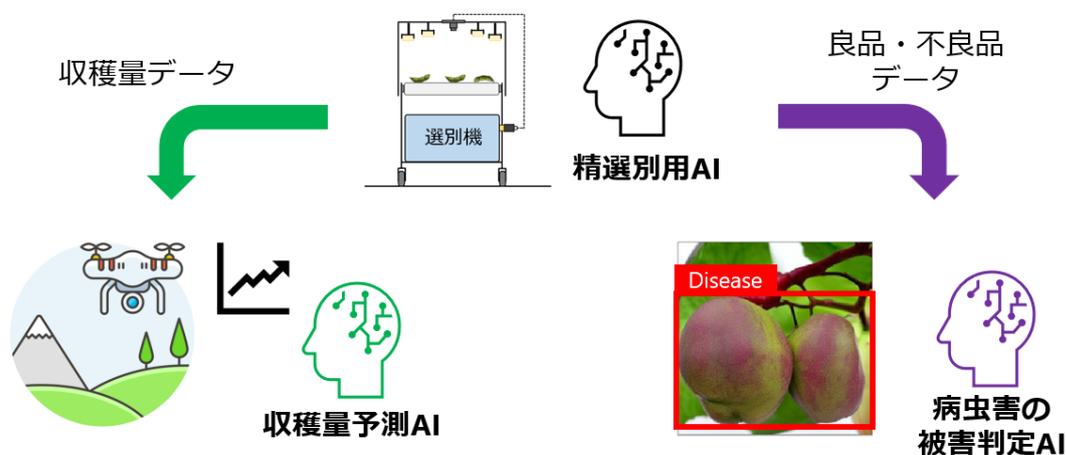


図 4-11 精選別用 AI が他の AI へデータをシェアして
それぞれの性能向上へ寄与するイメージ

この時、対象となるほ場のエダマメを精選別機に通すことで、正確な収穫量や良品・不良品の割合を算出できる。選別データを利用して収穫量予測 AI などを開発する際は、空撮から得られる圃場でのエダマメの生育データと選別から得られる収穫量などの正解データを正しく紐づけておく必要がある。この場合、生産者は、AI 開発を考慮して、対象となるほ場を精選別機に設定し、異なるほ場で収穫されたエダマメを混合して精選別機に通さないようにする必要がある。

選別情報のデータ収集と二次利用は、最終的に精選別機の普及促進に貢献する可能性がある。精選別機から得られる選別データは、別の業務や作業への二次利用が推進されることで、データインフラを構築する上で必要不可欠になる。昨今、スマートフォンが電話だけでなくカメラやインターネットのブラウジング、支払い機能などにも利用されている。今後 AI を搭載した選別機も、スマートフォンのように、本来の用途である選別以外の用途へ積極的に使われ、精選別機の導入が推奨されるようになると予想される。このような未来が到来することを考慮し、精選別用 AI の高精度化の研究は、選別精度だけでなく先述した他の業務や作業の精度向上にも繋がるため、継続して行う必要がある。

また、本研究の物体検出 AI を用いた外観選別の知見は、エダマメに限らず他の農作物にも応用可能であり、選別の省力化、高精度化、選別に関する二次利用に貢献する可能性がある。様々な選別対象について、選別精度向上や得られた選別データを活用する様々なアイデアを研究者だけでなく企業や生産者など幅広く情報共有することで、データに基づいた農作物生産の管理が加速することを期待する。

謝辞

本研究の遂行と論文をまとめるにあたり、山形大学農学部 of 片平光彦教授には、主指導教員として研究実施と論文作成の全般でご指導とご助言をいただきました。山形大学農学部 of 小林隆准教授には、生産現場の観点からご指導とご助言を頂きました。弘前大学 of 張樹槐教授には、画像処理の技術的観点からご指導とご助言をいただきました。

本研究の遂行にあたり、DL による AI 開発や AI の開発環境の整備、本論文の AI に関する議論や助言、添削全般に関して、山形大学農学部 of 市浦茂先生には多大なご指導をいただきました。また、エダマメのデータ収集には、株式会社ガオチャオエンジニアリング of 高橋史夫代表取締役及び株式会社石井製作所の石井智久代表取締役、庄司翔大様にご協力を頂きました。そして、Dhirendranath Singh 博士には論文作成にあたり英文校正や AI に関する議論を、山形大学農学部生産機械研究室の方々には研究のデータ整備や議論にご協力いただきました。

最後に、これまで私の大学生活を支えてくださった両親に心から感謝します。

末尾ながら、本研究が関係各位の多大なご支援によってここにあることを記して、改めまして心より感謝申し上げます。

引用文献

- 1) 藤田智, 2017. 恵泉野菜の文化史(13): エダマメ, 恵泉女学園大学園芸文化研究所報告, 13, 121-127.
- 2) Konovsky, J., Lumpkin, T. A., McClary, D., 2020. Edamame: the vegetable soybean. In *Understanding the Japanese Food and Agrimarket*, CRC press, 173-181.
- 3) 三宅紀子, 酒井清子, 五十嵐歩, 鈴木恵美子, 倉田忠男, 2007. ゆで加熱条件下におけるエダマメ中の呈味成分およびビタミン C 含量の変動, 日本調理科学会誌, 40, 3, 189-192.
- 4) 鈴木晶, 2011. ニューヨークとパリ, 異文化, (12), 45-51.
- 5) 下渡敏, 2003. 食品製造業のグローバリゼーションと国内原料調達, 農業経済研究, 75, 2, 47-54.
- 6) 独立行政法人農畜産業振興機, 2017. 海外情報: タイのえだまめの生産及び輸出動向,
https://vegetable.alic.go.jp/yasaijoho/kaigaijoho/1708_kaigaijoho02.html,
Accessed October. 17. 22.
- 7) Sharma, K. P., 2013. Varietal yield stability and appropriate management for quality organic Edamame production, Vernon, BC, Organic BC.
- 8) 農林水産省, 2020. 作物統計調査/作況調査(野菜) 確報 野菜生産出荷統計,
https://www.maff.go.jp/j/tokei/kouhyou/sakumotu/sakkyou_yasai/index.html,
Accessed July. 8. 2022.
- 9) 高柳長直, 2002. 輸入野菜増加傾向下における野菜産地の形成—山形県鶴岡市の枝豆産地の事例—, 農村研究, 94, 46-60.
- 10) 農林水産省, 2020. 農業物価統計調査 令和2年度農業物価統計,
<https://www.e-stat.go.jp/stat->

search/files?page=1&layout=datalist&toukei=00500204&tstat=000001014483&cycle=7&year=20200&month=0&tclass1=000001019836&tclass2=000001158368, Accessed July. 8. 2022.

- 11) 農作業学会, 農作業データ集, https://jsfwr.org/fw_data/index.html, Accessed July. 11. 2022.
- 12) 岩田隆, 白幡啓一, 1979. エダマメ収穫後の品質変化とその防止 (第1報) 品質変化に関係する要因とガス組成及び葉付き包装の効果, 園藝學會雜誌, 48, 1, 106-113.
- 13) 鶴沼秀樹, 片平光彦, 2007. エダマメ収穫調製機械作業体系の経済的導入規模, 東北農業研究, 60, 235-236.
- 14) 農林水産省, 2022. 農業労働力に関する統計, <https://www.maff.go.jp/j/tokei/sihyo/data/08.html>, Accessed July. 8. 2022.
- 15) 佐々木豊, 鈴木正肚, 陳山鵬, 齋藤友宏, 2000. マシンビジョンによる枝豆の知的選別機の開発—形状情報を利用した枝豆の等級選別の可能性—, 農業機械学会誌, 62, 6, 92-97.
- 16) 徳田勝, 山本博昭, 川村恒夫, 伊藤博通, 松山善之助, 2002. 特産ダイズ丹波黒枝豆の選別に関する研究 (第1報) —形状解析による粒数, 粒種, 湾曲莢の判別—, 農業機械学会誌, 64, 4, 68-75.
- 17) 片平光彦, 張樹槐, 大泉隆弘, 後藤恒義, 鶴沼秀樹, 田村晃, 後藤克典, 2011. 画像処理によるエダマメの選別方法に関する研究 (第2報) 試作エダマメ選別機の性能評価. 農業機械学会誌, 73(2), 127-134.
- 18) LeCun, Y., Bengio, Y., Hinton, G., 2015, Deep learning, Nature, 521, 436-444.
- 19) 浅井伸一, 江間有沙, 工藤郁子, 巢籠悠輔, 瀬谷啓介, 松井孝之, 松尾豊, 2018. 深層学習教科書 ディープラーニング G 検定 (ジェネラリスト) 公式テキスト (初版), 株式会社翔泳社, 東京.

- 20) 齋藤康毅, 2019, ゼロから作る DeepLearning—Python で学ぶディープラーニングの理論と実装 (初版), 日経印刷株式会社, 東京.
- 21) 福島邦彦, 1979. 位置ずれに影響されないパターン認識機構の神経回路モデル—ネオコグロナー, 電子情報通信学会論文誌, J62-A, 10, 658-665.
- 22) LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86, 11, 2278-2324.
- 23) Krizhevsky, A., Sutskever, I., Hinton, E. G., 2012. ImageNet Classification with Deep Convolutional Networks. Advances in Neural Information Processing System 25. NIPS2012.
- 24) Simonyan, K., Zisserman, A., 2015. VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, <https://arxiv.org/pdf/1409.1556.pdf>, Accessed July. 13. 2022.
- 25) Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2015. Rethinking the Inception Architecture for Computer Vision, <https://arxiv.org/pdf/1512.00567.pdf>, Accessed July. 13. 2022.
- 26) He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image recognition. <https://arxiv.org/pdf/1512.03385v1.pdf>, Accessed July. 13. 2022.
- 27) He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on Imagenet Classification. IEEE International Conference on Computer Vision (ICCV), 1026-1034.
- 28) Zhang, H., Cloutier, R., 2022. Review on One-Stage Object Detection Based on Deep Learning, EAI Endorsed Transactions on e-Learning, 7, 23, 1-10.

- 29)Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.
<https://arxiv.org/pdf/1506.01497.pdf>, Accessed July. 15, 2022.
- 30)Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., Berg, C, A., 2016. SSD:Single Shot MultiBox Detector.
<https://arxiv.org/pdf/1512.02325v5.pdf>. Accessed July. 15, 2022.
- 31)Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection.
<https://arxiv.org/pdf/1506.02640v5.pdf>, Accessed July. 15. 2022.
- 32)Tan, M., Pang, R., Le, Q., 2020. EfficientDet: Scalable and Efficient Object detection, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 10781-10790.
- 33)Bochkovskiy, A., Wang, C., Liao, H. 2020, YOLOv4: Optimal Speed and Accuracy of Object Detection, <https://doi.org/10.48550/arXiv.2004.10934>, Accessed July. 15. 2022.
- 34)上垣俊平, 荒木秀和, 戸島亮, 新崎誠, 上田大介, 山崎龍次, 2018. 環境認識に AI を用いたトマト収穫ロボット. *Panasonic Technical Journal*, 64(1), 54-59.
- 35)朝田諒, 飯田訓久, 村主勝彦, 増田良平, 2019. ロボットコンバインのためのステレオカメラを用いた物体検出および衝突回避. *農業食料工学会誌*, 81(6), 392-402.
- 36)Apolo, O., Martinez-Guanter, J., Egea, G., Raja, P., Perez-Ruiz, M., 2020. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV, *European Journal of Agronomy*, 115, 126030.
- 37)Unay, D., 2022. Deep learning based automatic grading of bi-colored apples using multispectral images, *Multimedia Tools and Applications*.

- 38)Chen, Y., An, X., Gao, S., Li, S., Kang, H., 2021. A Deep Learning-Based Vision System Combining Detection and Tracking for Fast On-Line Citrus Sorting, *frontiers in Plant Science*, 12, 622062.
- 39)Bhupendra, Moses, K., Miglani, A., Kankar, P., 2022. Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset, *Computers and Electronics in Agriculture*, 195, 106811.
- 40)Tan, M., Le, Q., 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, <https://arxiv.org/abs/1905.11946>, Accessed July. 15. 2022.
- 41)岡谷貴之, 2014. ディープラーニング, *映像情報メディア学会誌*, 68, 6, 466-471.
- 42)Cortes, C., Vapnik, V., 1995. Support-vector Networks, *Machine Learning*, 20, 273-293.
- 43)Aggelopoulou, A., Bochtis, D., Fountas, S., Swain, K., Gemtos, T., Nanos, C. 2011. Yield prediction in apple orchards based on image processing, *Precision Agric*, 12, 448-456.
- 44)Lin, G., Tang, Y., Zou, X., Li, J., Xiong, j., 2019. In-field citrus detection and localization based on RGB-D image analysis. *Biosystems Engineering*, 186, 34-44.
- 45)Xie, S., Girshick, R., Dollar, P., Tu, Z., He, K., 2017. Aggregated Residual Transformations for Deep Neural Networks, <https://arxiv.org/pdf/1611.05431.pdf>, Accessed August. 10. 2022.
- 46)中山英樹, 2015. 深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習, *信学技報*, 115, 146, 55-59.
- 47)Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P., Zitnick, C., 2012. Microsoft COCO: Common Objects in Context,

ECCV 2014:

European Conference on Computer Vision, 740-755.

- 48)鄭 宏杰, 中林 靖, 増田 正人, 西 宏起, 山中 大介, 高橋 伸一郎, 伯野 史彦, 宮内 陽介, 岡崎 敬, 吉田 一寛, 塩谷 隆二, 2020. CNNに基づく鶏肝臓画像における脂肪肝分類に対する転移学習の効果検証, 日本計
算工学会論文集, 1, 20201003.
- 49)Zoph, B., Cubuk, E., Ghiasi, G., Lin, T., Shlens, J., Le, Q., 2020. Learning
Data Augmentation Strategies for Object Detection,
<https://arxiv.org/pdf/1906.11172.pdf>, Accessed August 2. 2022.
- 50)Suchet, B., James, U., 2017. Deep Fruit Detection in Orchards.
<https://arxiv.org/pdf/1610.03677.pdf>, Accessed August. 2. 2022.
- 51)山口永, 2021. MLOps による機械学習プロセスの高速化と継続的にサ
ービス価値を提供するための仕組みづくり, NTT 技術ジャーナル,
2021, 3, 69-75.
- 52)中田典生, 2017. 画像診断における人工知能活用の現状と展望, 日腎
会誌, 59, 7, 1064-1070.
- 53)Hughes, D., Salathé, M., 2015. An open access repository of images on
plant health to enable the development of mobile disease diagnostics,
<https://arxiv.org/abs/1511.08060>, Accessed October. 17. 22.
- 54)Ferentinos, K. P., 2018. Deep learning models for plant disease detection
and diagnosis, Computers and electronics in agriculture, 145, 311-318.
- 55)彌富仁, 2019. 植物病害自動診断技術の動向と課題, 日本神経回路学
会誌, 26, 4, 123-134.
- 56)森智洋, 市浦茂, 片平光彦, 2021. ディープラーニングを用いたエダ
マメの外観精選別用 AI の作成と精度評価 (第 1 報) 物体検出 AI に
よるエダマメの外観精選別の有効性の検討. 農業食料工学会誌,

- 83(3), 163-171.
- 57)Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement.
<https://arxiv.org/pdf/1804.02767v1.pdf>. Accessed July. 21. 2022.
- 58)Sultana, F., Sufian, A., Dutta, P., 2020. A Review of Object Detection Models based on Convolutional Neural Network,
<https://arxiv.org/pdf/1905.01614.pdf>, Accessed July. 21. 2022.
- 59)Dai, J., Li, Y., He, H., Sun, J., 2016. R-FCN: Object Detection via Region-based Fully Convolutional Networks, <https://arxiv.org/pdf/1605.06409.pdf>, Accessed July. 21. 2022.
- 60)Koirala, A., Walsh, K. B., Wang, Z., McCarthy, C., 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of ‘MangoYOLO’. Precision Agriculture, 20, 1107-1135.
- 61)野村泰稔, 村尾彩希, 坂口幸広, 古田均, 2017. 深層畳み込みニューラルネットワークに基づくコンクリート表面のひび割れ検出システム. 土木学会論文集, 73(2), 189-198.
- 62)Tsung-Yi, L., Piotr, D., Ross, G., Kaiming, H., Bharath, H., Serge, B., 2017. Feature Pyramid Networks for Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2117-2125.
- 63)DMP BLOG, 2018. A Closer Look at YOLOv3,
<https://blog.dmprof.com/post/a-closer-look-at-yolov3>, Accessed August. 4. 2022.
- 64)農業機械学会, 1996. 生物生産機械ハンドブック (初版), 株式会社コロナ社, 東京, 813-815.
- 65)農文協, 2004. 野菜園芸大百科 (第2版) 8 エンドウ・インゲン・ソラマメ・エダマメ・その他マメ. 社会法人農山漁村文化協会, 東

- 京, 383-465.
- 66)阿部利徳, 2011. エダマメにおけるダダチャマメ系品種の生育および成分特性, 育種学研究, 13, 1-10.
- 67)新潟県, えだまめの出荷規格,
<https://www.pref.niigata.lg.jp/uploaded/attachment/145947.pdf>, Accessed September. 21. 2022.
- 68)原聡, 2019. 説明可能な AI. 人工知能, 34, 4, 577-582.
- 69)Ramprasaath, S., Michael, C., Abhishek, D., Ramakrishna, V., Devi, P., Dhruv, B., 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, IEEE International Conference on Computer Vision (ICCV), 618-626.
- 70)Haofan, W., Zifan, W., Mengman, D., Fan, Y., Zijian, Z., Sirui, D., Piotr, M., Xia, H., 2020. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 24-25.
- 71)西尾真由子, 栗栖雄一, 2020. 橋梁点検部材損傷度判定 CNN の可視化による判断根拠の理解と活用, AI・データサイエンス論文集, 1, J1, 92-99.
- 72)Yebasse, M., Shimelis, B., Warku, H., Ko, J., Cheoi, K. J., 2021. Coffee Disease Visualization and Classification, Plants, 10, 6, 1257.
- 73)Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., Jagersand, M., 2020. U2-Net: Going Deeper with Nested U-structure for Salient Object Detection, Pattern Recognition, 106, 107404.
- 74)岡谷貴之, 2016. 画像認識のための深層学習の研究動向: 畳込みニューラルネットワークとその利用法の発展, 人工知能, 31, 2, 169-179.
- 75)Rizwan Iqbal, H. M., Hakim, A, 2022. Classification and grading of

- harvested mangoes using convolutional neural network. International Journal of Fruit Science, 22, 1, 95-109.
- 76) Zoph, B., Vasudevan, V., Shlens, J., Le, Q. V., 2018. Learning transferable architectures for scalable image recognition, In Proceedings of the IEEE conference on computer vision and pattern recognition, 8697-8710.
- 77) 独立行政法人農畜産業振興機構, 2019. 野菜ブック, https://www.alic.go.jp/y-suishin/yajukyu01_000313.html, Accessed December. 4. 2022.
- 78) Northcutt, C. G., Athalye, A., Mueller, J., 2021. Pervasive label errors in test sets destabilize machine learning benchmarks, <https://arxiv.org/abs/2103.14749>, Accessed September. 23. 2022.
- 79) Krizhevsky, A., Hinton, G., 2009. Learning multiple layers of features from tiny images, <http://www.cs.utoronto.ca/~kriz/learning-features-2009-TR.pdf>, Accessed September. 23. 22.
- 80) Toshiba Clip, 「東芝の AI, 作って終わりから次のステージへ ～AI の学習サイクルを自動化する「MLOps」構想」, <https://www.toshiba-clip.com/detail/p=7077>, Accessed September. 23. 2022.
- 81) Liu, J., Wang, X., 2021. Plant diseases and pests detection based on deep learning: a review, Plant Methods, 17, 1, 1-18.
- 82) Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M., 2021. Scaled-yolov4: Scaling cross stage partial network, In Proceedings of the IEEE/cvf conference on computer vision and pattern recognition, 13029-13038.
- 83) Chien-YAO, W., I-Hau, Y., Hong-Yuan, M. L., 2021. You Only Learn One Representation: Unified Network for Multiple Tasks, arXiv, <https://arxiv.org/pdf/2105.04206v1.pdf>, Accessed September. 29. 2022.

- 84) Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M., 2022. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv, <https://arxiv.org/abs/2207.02696>, Accessed September. 29. 2022.
- 85) Wojke, N., Bewley, A., Paulus, D., 2017. Simple online and realtime tracking with a deep association metric, In 2017 IEEE international conference on image processing (ICIP), 3645-3649).
- 86) Mikel Broström, 2022. Real-time multi-camera multi-object tracker using YOLOv5 and StrongSORT with OSNet, https://github.com/mikel-brostrom/Yolov5_StrongSORT_OSNet, Accessed October. 5. 2022.
- 87) Joseph, L., YOLO: Real-Time Object Detection. <https://pjreddie.com/darknet/yolo/>, Accessed November 7. 2022.
- 88) 九州農政局, 食品企業課, 2017. 農産物の物流について, <https://www.tb.mlit.go.jp/kyushu/content/000091492.pdf>, Accessed October. 5. 2022.
- 89) 伊津野範博, 2021. 地方運送業の労働力確保に関する研究: 熊本県トラック協会による労働力確保実態調査より, 熊本学園商学論集, 25, 2, 107-140.
- 90) 株式会社ガオチャオエンジニアリング, <https://gaoqiao-eng.com/pit-eds-mini01>, Accessed November. 7.
- 91) 外務省, SDGs とは?, <https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/about/index.html>, Accessed November. 16. 2022.
- 92) 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, 3, 224-235.

- 93)奥村俊夫, 2021. リモートセンシングにおける人工知能 (AI) 活用, 日本リモートセンシング学会誌, 41, 2, 315-318.
- 94)Dharani, M. K., Thamilselvan, R., Natesan, P., Kalaivaani, P. C. D., Santhoshkumar, S., 2021. Review on crop prediction using deep learning techniques. In Journal of Physics: Conference Series, 1767, 1, 012026.