

	刊 トモロ
氏 名	森 智洋
本籍（国籍）	石川県
学位の種類	博士（農学）
学位記番号	連研第 840 号
学位授与年月日	令和 5 年 3 月 2 3 日
学位授与の要件	学位規則第 5 条第 1 項該当課程博士
研究科及び専攻	連合農学研究科 地域環境創生学専攻
学位論文題目	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)
学位審査委員	主査 山形大学教授 片平 光彦 副査 山形大学准教授 小林 隆 副査 弘前大学教授 張 樹槐 副査 岩手大学教授 小出 章二

### 論 文 の 内 容 の 要 旨

本研究は、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を目的とし、ディープラーニング（以下、DL）で高精度なエダマメ外観精選別（以下、精選別）用の物体検出 AI を開発するための学習データの構成について検討した。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い、学習データに起因する精度低下の要因を絞り込んだ。次に、AI の分類根拠可視化技術「Score-CAM」で AI の精度低下因子を特定し、学習データの観点から精度を改善する方法を考察した。最終的に、本研究で得られた知見から、エダマメ精選別用の MLOps を考案した。

1 つ目の実験では、学習データのどのような内容が AI の精度変化に影響するか絞り込むため、学習データに含むエダマメ品種の違いが物体検出 AI の精度に与える影響を調査した。品種に注目した理由は、先行研究で学習データの画像に写る品種が変化すると物体検出 AI の精度が変化し、品種が AI の精度変化の因子になると考えたためである。物体検出 AI に分類させる項目は、良品 1 項目と不良品 6 項目（一粒、欠粒、奇形、未熟、変色、さや割れ）の計 7 項目を設定した。エダマメ画像は庄内三号と庄内一号、グリーン 75 の 3 品種を使用し、品種の組み合わせを変えた 7 つの学習データセットを設定した。DL で開発した物体検出 AI は、YOLOv3 と Faster R-CNN の 2 種類である。精度評価は庄内三号の画像を使用して、適合率と再現率、F 値で評価した。適合率と再現率、F 値の平均は、YOLOv3 と Faster R-CNN の両方で選別対象品種の画像を学習データに加えた物体検出 AI の方が、他品種の画像のみを学習データに含む物体検出 AI よりも有意に高くなった。F 値の平均はいずれの物体検出 AI 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、学習データの品種の違いが物体検出 AI の精度低下因子になることが分かった。

次に、品種の異なる学習データで開発した画像認識 AI へ Score-CAM を適用した。これは、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、その因子の影響を考慮した学習データの構築が可能になるためである。2 つ目の実験では、良品 1 項目と不良品 6 項目の計 7 項目を分類できる ResNet152 と VGG16 の画像認識 AI を開発した。使用した品種は庄内三号と庄内一号、秘伝の 3 品種であり、これらの 3 品種で 3 つの学習データセットを設定した。DL で品種ごとの画像認識 AI を開発した後、Score-CAM で細分類項目の分類根拠を特定した。細分類項目は、画像認識 AI に分類させる 7 項目をさらに 19 項目へ細分化した項目である。テスト画像は、品種・細分類項目ごとに 30 枚ずつ用意して、学習データの品種と同じ品種のテスト画像を画像認識 AI へ入力した。画像認識 AI の各細分類項目の分類根拠は、Score-CAM のヒートマップの発生パターンが最多の部分とした。学習データの品種の違いで分類根拠が異なった細分類項目の数は、ResNet152 で 15 項目、VGG16 が 12 項目であった。各分類項目の特徴を学習済みの AI は、品種に関わらず分類根拠が一致すると考えられるが、ResNet152 と VGG16 で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外観情報も特徴量として利用している。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（品種バイアス）。また、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は、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ため、複数の品種の画像を組み合わせることによって各外観品質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り、DL による品種バイアスの解消につながる。

本研究から、学習データに含まれる AI の精度低下の特定は、精度の評価を通じて因子を推測し、その条件を変えた AI に CNN の分類根拠可視化技術を使って行う。学習データに関する AI の精度低下は、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て解決する。

物体検出 AI が搭載された精選別機は、選別量や良品・不良品の割合などを数値化できるため、異業種とのデータ連携による集荷の最適化や選別項目の決定、他の AI での正解データとしての利用など精選別以外にも活用できる。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、選別の省力化や高精度化に貢献できる。

## 論文審査の結果の要旨

本研究は、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を目的とし、ディープラーニング（以下、DL）で高精度なエダマメ外観精選別（以下、精選別）用の物体検出 AI を開発するための学習データの構成について検討した。

1 つ目の実験では、学習データに含むエダマメ品種の違いが物体検出 AI の精度に与える影響を調査した。物体検出 AI に分類させる項目は、良品 1 項目と不良品 6 項目の計 7 項目、エダマメ 3 品種を組み合わせた 7 つの学習データセットを設定した。DL で開発した物体検出 AI は、YOLOv3 と Faster R-CNN の 2 種類である。精度評価は適合率と再現率、F 値で行った。適合率と再現率、F 値の平均は、いずれも選別対象品種の画像を学習データに加えた物体検出 AI で他品種の画像のみを学習データに含む物体検出 AI よりも高くなった。F 値の平均はいずれの物体検出 AI も学習データに選別対象の品種が含まれていないと低下した。

2 つ目の実験では、良品 1 項目と不良品 6 項目の計 7 項目を分類できる ResNet152 と

VGG16の画像認識AIを開発した。学習データセットは市販のエダマメ3品種で設定した。DLで品種ごとに開発した画像認識AIは、Score-CAMのヒートマップを用いて細分類項目の分類根拠を特定した。学習データの品種の違いで分類根拠が異なった細分類項目の数は、ResNet152で15項目、VGG16が12項目であった。ResNet152とVGG16で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。エダマメの精選別用AIでは、各外観品質の情報の他に品種固有の外観情報も特徴量として利用しており、それが精度低下の因子になる。また、物体検出AIでは、複数の品種の画像を組み合わせることで外観品質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り、DLによる精度低下の解消になる。

本技術は今後農業界の進展に寄与するスマート農業に関する技術研究であるため、研究者のみならず農業事業者が利用できる基礎的知見として極めて価値が高いといえる。よって、本審査委員会では、「岩手大学大学院連合農学研究科博士学位論文審査基準」に則り審査した結果、本論文を博士（農学）の学位論文として十分価値のあるものと認めた。

#### 学位論文の基礎となる学術論文

##### 1. 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

##### 2. 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2報）－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の外観精選別精度に与える影響－

農業食料工学会誌, 83(3) : 172-181