

〔論 文〕

機械学習によるキュウリ果実の外観品質評価

八坂 亮祐^{*1}・辻 陽仁^{*2}・今村 彩菜^{*3}・岡 美琴^{*2}・比嘉 夏辰^{*2}平田 一翔^{*2}・伊福 佑耶^{*2}・小田 まり子^{*2,4}

Evaluation of Appearance Quality of Cucumber Fruits Using Machine Learning

Ryosuke YASAKA^{*1}, Haruhito TSUJI^{*2}, Ayana IMAMURA^{*3}, Mikoto OKA^{*2}, Keitatsu HIGA^{*2},Kazuto HIRATA^{*2}, Yuya IFUKU^{*2} and Mariko ODA^{*2,4}

Abstract

Amidst the surge in AI technological advancements, the application of computer vision for assessing the appearance quality of fruits and vegetables has garnered significant interest. This study endeavors to enhance the consistency, efficiency, and objectivity in quality assessment by employing image recognition AI to evaluate the appearance of cucumber fruits. We prepared a dataset comprising images of fruits categorized into three quality grades—A, B, and C—based on fruit distortion and skin condition. The research involved developing a detection model using YOLOv8 and a quality grade classification model employing a convolutional neural network (CNN). Furthermore, we constructed a model to estimate fruit length by comparing the pixel count of the detected fruit with a horticultural label serving as a length standard. Testing these models yielded several key findings: (1) the object recognition model demonstrated high accuracy in detecting the position of the fruit within the image and its detection range; (2) the transferred EfficientNet CNN model exhibited the highest accuracy in classifying fruit grades; (3) the developed method enabled estimating the fruit length with an accuracy suitable for production site applications. Future efforts will focus on expanding the image dataset, minimizing labeling errors, and enhancing classification accuracy by incorporating multiple quality classification criteria in the labeling process.

Key Words : Machine Learning, Object Detection, Convolutional Neural Network, Smart Agriculture Technology, Data-driven Agriculture

1. 緒 言

現在の日本では、農業従事者の高齢化、後継者不足など深刻な課題に直面している。自営農業従事者の人口は平成 22 年から令和 4 年にかけて 205 万人から 123 万人となり、12 年間で約 40%減少、65 歳以上の高齢者が占める割合は 70% に達している^{(1),(2)}。新規就農者数の減少も顕著であり、後継者不足により技術継承が難しく、日本の農業技術の低下が懸念されている。そのため新規就農者を増やす方策が重要となるが、併せて安定した生産量と優れた品質の農作物を生産するための技術開発も急務となる。

このような状況下において、情報技術、特に人工知能 (AI) の進展により、農業は大きな変革期を迎えている。AI の活用は、収穫量の予測、病害の早期発見、効率的な育成管理など、多岐にわたる領域でその可能性を示している⁽³⁾。特に農作物の品質管理と収穫量の向上に関する研究が盛んになされている。例えば、ダイズの遺伝的特性と生育環境のデータを用いて機械学習モデルを構築し、最適な品種改良のための品種選択や交配戦略を改善する研究⁽⁴⁾ などが行われている。

この技術革新の流れの中で、果菜類の外観品質評価は注目される領域の 1 つである。一般に果菜類の外観品質は、果形、色彩、果皮部の状態および果長などの物理的特徴に基づいて評価される。これらの特徴を正確かつ迅速に把握することは、品質管理の効率化に不可欠であるが、現行の評価方法は、多くの場合人間による目視観察に依存しているため、一貫性の

^{*1} PC サポートセンター、^{*2} 情報ネットワーク工学科、^{*3} 建築・設備工学科、^{*4} AI 応用研究所
令和 6 年 1 月 10 日受理

欠如、主観的な判断の介入、作業労力の増加といった問題が生じている．ここで AI による画像解析技術を用いれば、評価者の経験や勘に依存しない客観的な評価が可能となる．さらには定量的な評価も可能となり、非侵襲的な評価方法であるためデリケートな果実を傷めるリスクも抑えられる．

そこで本研究では、果菜類の 1 つであるキュウリについて、従来の人間による評価手法に代わり、画像認識 AI を導入することで果実の外観特徴を正確に識別し、評価プロセスを一貫性のあるものにするとともに、品質評価の効率化と客観性の向上を目的とする．

2. キュウリ果実画像データセットの準備

本研究では、果実の商品価値を測る上で重要となる果実形質および果長を評価対象項目とした．果実の画像データセットは、果形の歪みや果皮部の状態に基づき等級分けを行い、品質の良い順に A, B および C の 3 等級でラベル付けを行った．等級分けを行った果実の例を果実形質とともに図 1(a) に示す．果実の等級分けは、株式会社久留米原種育成会のキュウリ果実評価基準 (図 1(b) および(c)) を用いた．まず、図 1(b) に示す部位ごとの形質により果実を 3 等級に振り分けた．果実形質の「曲がり」は商品価値に最も影響を与えるため、「曲がり」がある果実については、図 1(c) に示す隙間を測定し、測定結果と果実形質の評価を踏まえた上で最終的な等級を決定した．等級分けを行った結果、画像データ数は果実 (A, B および C 級) がそれぞれ 224, 286 および 118 枚となった．また、果長測定の長さ基準として用いるため、園芸用ラベル (縦幅約 8.8cm) の画像データを 197 枚用意した．

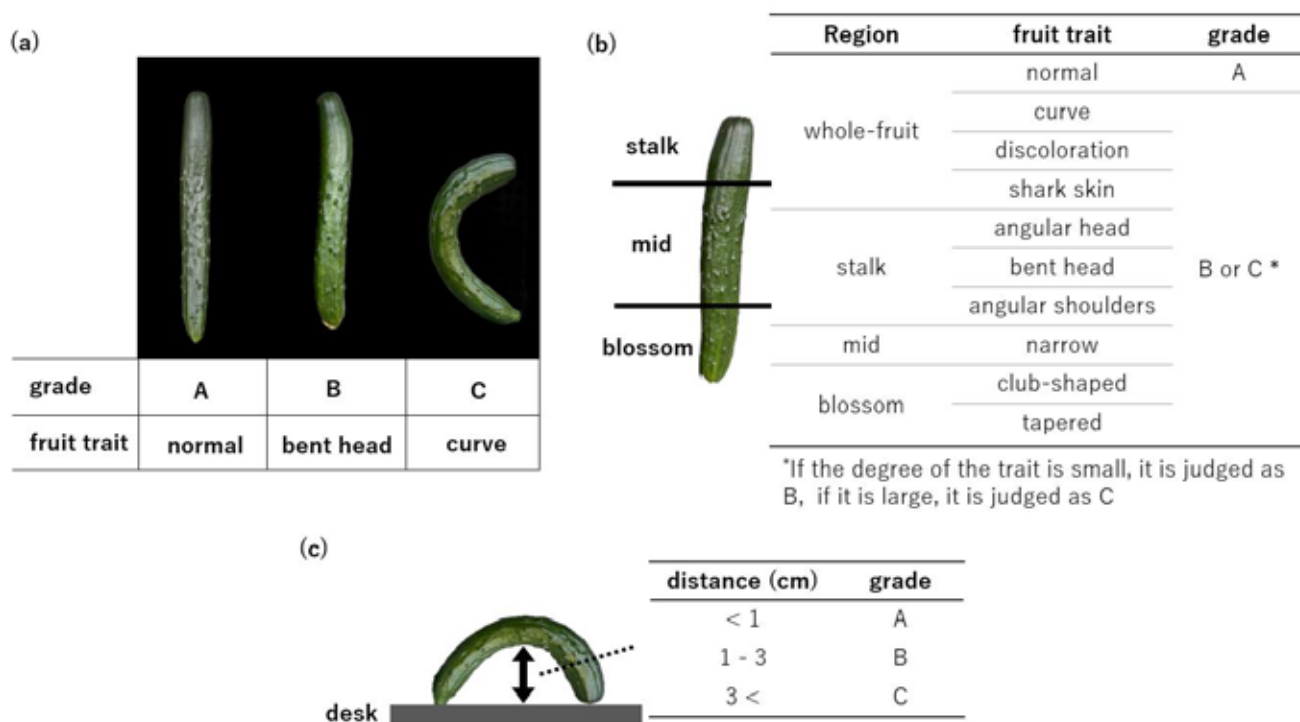


Fig. 1 Cucumber fruit grading criteria used in this study

3. 物体認識モデルの構築

3・1 物体認識モデルの構築手順

画像上の果実および園芸用ラベルを認識する物体認識モデルの構築を試みた．モデル構築の手順は次の通りである．

- (1) 果実および園芸用ラベルの画像に対して、Roboflow⁽⁵⁾ を用いてアノテーション付けを行った (図 2)．今回はインスタンスセグメンテーションによる 4 クラス (果実 A, B および C 級、園芸用ラベル) 分類を目的としたアノテーション付けを行った．

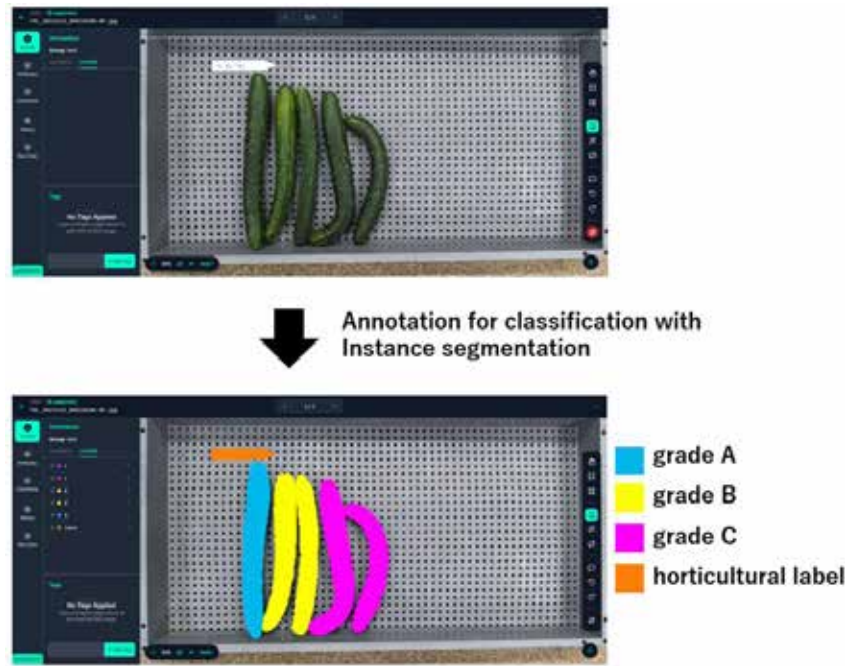


Fig. 2 Roboflow usage screen used for annotation

- (2) アノテーション付けを行った画像に対してデータ数を3倍とする水増し処理を行った。処理方法としては、過学習抑制のためのノイズ処理、および物体検出位置や検出範囲の精度向上を図るため、ぼかし処理を採用した。
- (3) 物体認識には物体認識アルゴリズム YOLO の最新版 (2023 年 12 月現在) である YOLOv8⁽⁶⁾ を用い、学習済みモデル (yolov8n-seg) の重みを初期値としてアノテーション画像による転移学習を行った。なお、訓練データと検証データの比率は 8 対 2 とし、学習時の epoch 数は 300、バッチサイズは 8、学習率は 0.01 とした。

3・2 物体認識モデルの性能評価

構築した物体認識モデルの性能評価を行った。指標としては Box Loss, Class Loss および Object Loss を用いて、検出位置 (バウンディングボックス) の正確さ、クラス分類の精度および検出範囲内の物体の有無の精度を評価した。各指標の値は損失関数で算出された正解値と予測値の誤差で与えられるため、学習ではこれらの値の最小化を目指した。モデルの転移学習時における Epoch ごとの各評価指標を図 3 に示す。図 3 では、Box Loss と Object Loss は学習が進むにつれて減少しているため、精度を高めながら学習ができていることがわかるが、一方 Class Loss については学習の早い段階で値が 1 程度で高止まりした。以上から、今回構築した物体認識モデルは物体位置の検出、検出範囲の精度は高められたが、クラス分類精度は低いと考えられた。

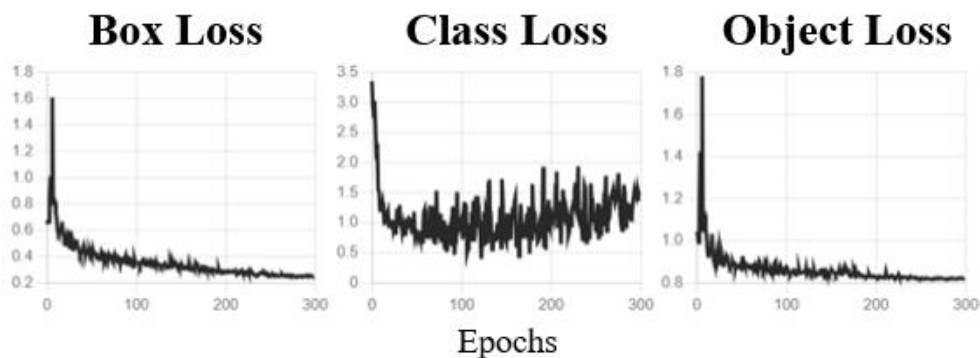


Fig. 3 Evaluation indexes of object recognition model

そこで、果実 A、B および C 級の画像をそれぞれ 67、86 および 25 枚含むテストデータを用いて、構築したモデルで 3 クラス分類の予測を行った。予測結果をまとめた混同行列を表 1 に示す。

Table 1 Confusion matrix of 3-class classification predicted using object recognition model

		predicted class		
		grade A	grade B	grade C
actual class	grade A	10	46	11
	grade B	40	33	13
	grade C	4	7	14

クラス分類精度を評価するために各クラスの正解率 (accuracy)、適合率と再現率の調和平均である F_1 値および 3 クラス分類全体の F_1 値のマクロ平均をそれぞれ式(1)、(2) および(3)で算出した。

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$F_1score = \frac{2TP}{2TP + FP + FN} \quad (2)$$

$$macroF_1 = \frac{1}{n} \sum_{i=1}^n (F_1score)_i \quad (3)$$

ここで、 n はクラス数、 TP 、 FP 、 FN および TN はそれぞれ真陽性、偽陽性、偽陰性および真陰性の予測数である。これらの式を用いて算出した値を表 2 に示す。クラスごとの正解率は C 級が約 80.3%で最も高く、B 級が約 40.4%と最も低かった。また、 F_1 値は C 級が約 44.4%で最も高く、A 級が約 16.5%と最も低かった。3 クラス分類全体での F_1 値のマクロ平均は約 33.1% であり、クラス分類精度は良いとは判断できず、図 3 で示した学習時の Class Loss の高止まりと矛盾のない予測結果となった。

Table 2 Evaluating of classification accuracy of object recognition model

class	accuracy	F_1 score	macro F_1
grade A	0.433	0.165	
grade B	0.404	0.384	0.331
grade C	0.803	0.444	

4. 畳み込みニューラルネットワークモデルの構築

前章で述べた物体認識モデルでは果実 3 等級のクラス分類は難しいと判断し、画像分類タスクに対して有用な畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) でのクラス分類を試みた。

4・1 画像データセットの準備

CNN モデルを用いて果実のクラスを予測する際、入力画像には分類対象である果実 1 本のみが含まれていることが好ましい。そこで、物体認識モデルを用いて検出した果実が画像内に占める領域を算出し、画像からそれら領域部分を切り出すこととした。具体的には、画像内において矩形領域で示される検出範囲の四隅の座標値 (x および y 座標) を算出し、当該範囲を切り出した。その際、より高い検出精度を目指すため、前章で述べた手順で新たに果実と園芸用ラベルの 2 クラスでアノテーション付けを行った画像で再学習した物体認識モデルを構築した。次に、データ拡張のため切り出した果

実画像データの水増し処理を行った。画像の左右反転および画像を -20° から 20° の範囲においてランダムに回転させる処理を行い、データ数を 3 倍に拡張させた。

4・2 モデルの構築

CNN は畳み込み層とプーリング層の 2 種類の層を含む順伝播型ネットワークである。本研究で構築した CNN モデルの構造を図 4 に示す。モデル構築には、Python の深層学習ライブラリである Keras を用いた。畳み込み層では入力画像に対してカーネルを適用し、畳み込み演算を行うことで画像の特徴抽出を行う。プーリング層ではダウンサンプリングを行うことで演算量の削減や特徴の移動不変性の付与を行う。また、本モデルでは入力画像サイズは $50 \times 50 \times 3$ (RGB)、過学習の抑制のため Dropout および L2 正則化を行い、各畳み込み層で用いる活性化関数には ReLU、出力層の活性化関数には Softmax を採用した。

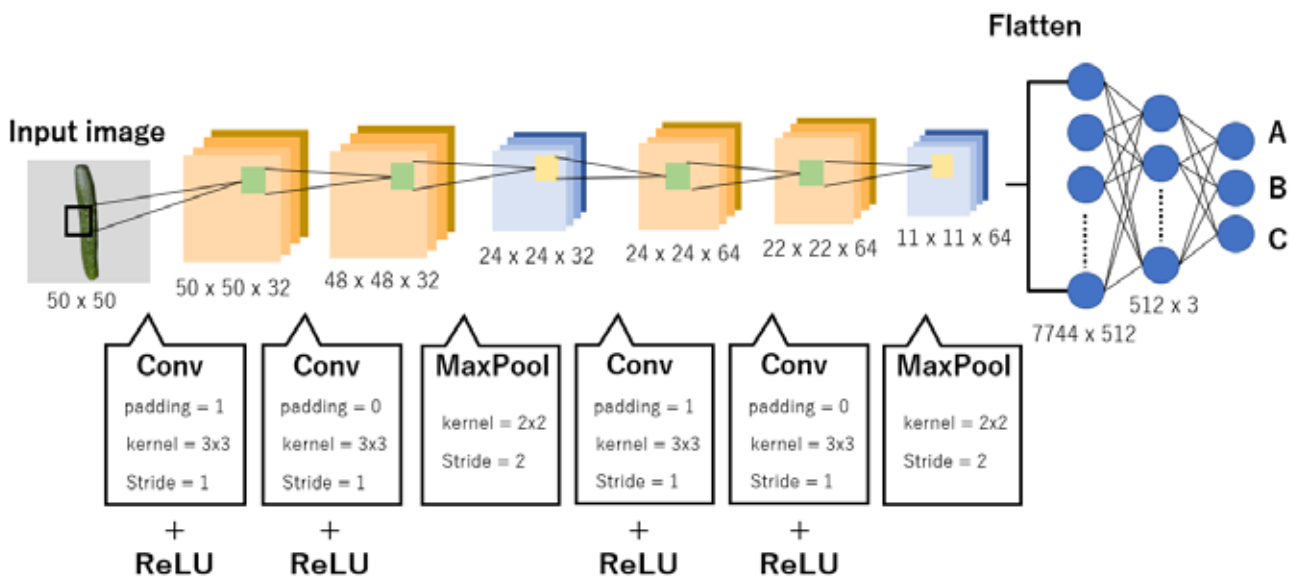


Fig. 4 Structure of the constructed convolutional neural network model in this study

4・3 モデルの学習

本 CNN モデルの学習では epoch 数を 100、バッチサイズを 8、学習率を 0.001 とし、最適化アルゴリズムとして Adam を採用した。また、訓練データと検証データの比率は 8 対 2 で分割し、ホールドアウト検証を行った。モデルの汎化性能を評価するための損失関数は多クラス交差エントロピー誤差 (CCE: Categorical Cross-Entropy) を用いた。モデルの学習過程で得られた epoch ごとの正解率と多クラス交差エントロピー誤差損失 (CCEL: Categorical Cross-Entropy Loss) の推移を図 5 に示す。学習が進むにつれて、訓練データの正解率は 100% に、CCEL は 0 に漸近した一方で、検証データの正解率は下がっていき、CCEL は上昇して収束しなかった。このことから、この一連の学習では顕著な過学習が起きていると考えられた。

そこで学習済みモデルによる転移学習により過学習の抑制を試みた。学習済みモデルには、大規模画像データセット ImageNet で学習済みの 16 層 CNN モデル VGG16⁽⁷⁾ および効率的な画像スケーリングによる高精度画像分類が可能な EfficientNet⁽⁸⁾ を用いることとした。EfficientNet については、最も軽量の EfficientNet-B0 を用いた。これら 2 種類の学習済みモデルの転移学習では、全結合層のみ今回のデータを用いて再構築し、学習済みモデルの重みは更新せずに学習を行った。両モデルの学習過程で得られた epoch ごとの正解率と CCEL の推移を図 6 に示す。両モデルともに当初の CNN モデルと同様、学習が進むにつれて、訓練データの正解率は 100% に、CCEL は 0 に漸近した。また、正解率については両モデルともに収束していないが、当初の CNN モデルの結果と比較すると、おおむね 2 倍以上の値を推移した。さらに転移学習した EfficientNet における CCEL に着目すると、転移学習した VGG16 における CCEL の 0.05% 程度を推移した。従って、転移学習した EfficientNet も過学習の傾向は見られるが、他モデルと比較して汎化性能が高い結果になったため、テストデータによる分類精度検証は本モデルを用いて行うこととした。

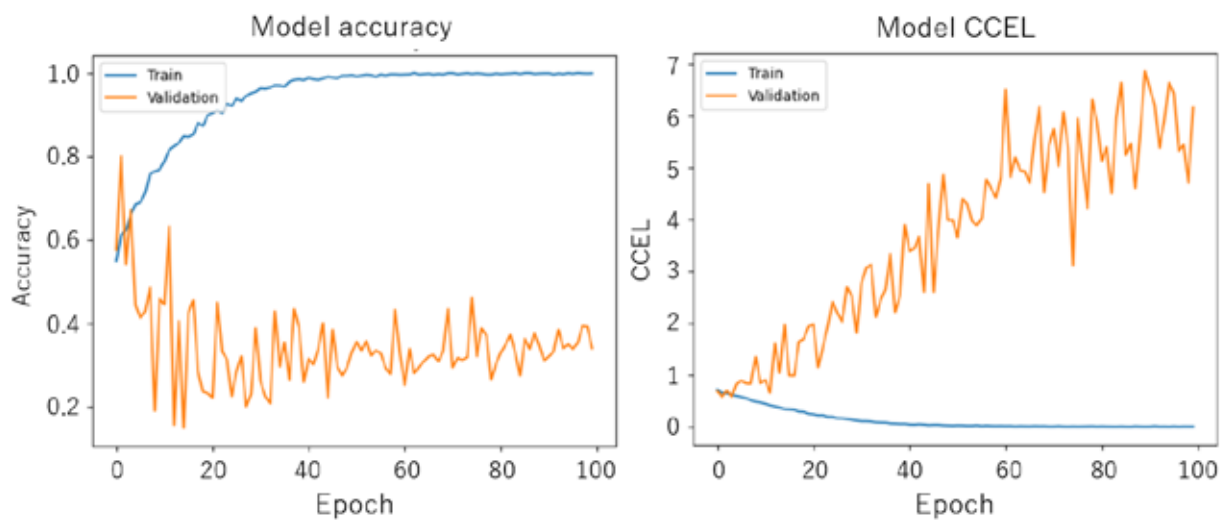
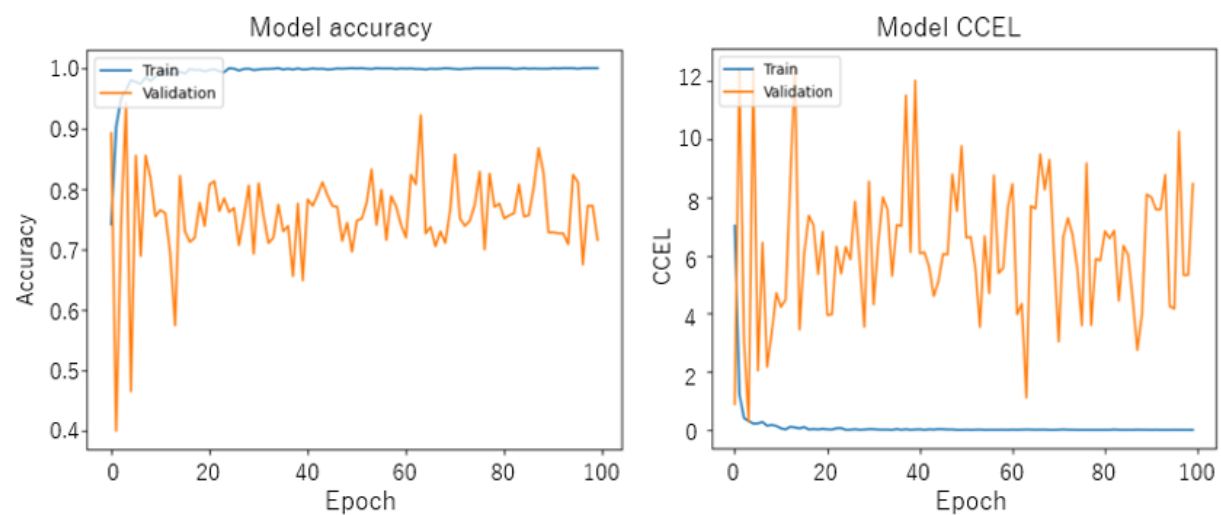
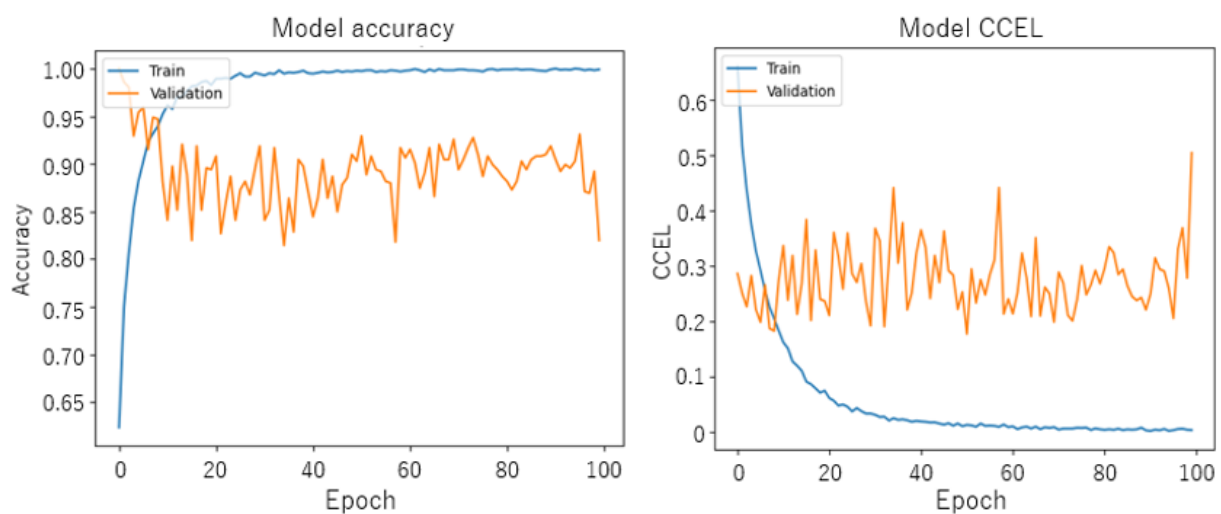


Fig. 5 Accuracy and the value of the loss function of CNN model during the learning process



(a) Transferred VGG16 model



(b) Transferred EfficientNet model

Fig. 6 Accuracy and the value of the loss function of transferred CNN models during the learning process

4・4 転移学習した EfficientNet の分類精度検証

転移学習した EfficientNet の分類精度検証には、物体認識モデルによる分類精度検証時に用いたテストデータと同一のものをを用いた。検証結果をまとめた混同行列を表 3 に示す。物体認識モデルによる予測 (表 1) との比較のため、両モデルの予測結果における混同行列のヒートマップを図 7 に示す。転移学習した EfficientNet では明らかに真陽性率が高くなったことがわかる。また、両モデルともに等級 A を B、または B を A と誤って分類していることが多く、比較的等級 C の分類精度が高かった。

Table 3 Confusion matrix for 3-class classification prediction using transferred EfficientNet

		predicted class		
		grade A	grade B	grade C
actual class	grade A	39	20	8
	grade B	31	46	9
	grade C	3	3	19



Fig. 7 Comparison of confusion matrix for 3-class classification prediction using object recognition model and transferred EfficientNet

次に、各クラスの正解率 (accuracy)、 F_1 値および 3 クラス分類全体の F_1 値のマクロ平均をそれぞれ式(1)、(2) および (3)で算出した値を表 4 に示す。ここで物体認識モデルによる分類精度評価指標 (表 2) と比較すると、すべての指標の値が高くなった。従って、果実 3 等級クラス分類におけるモデルの汎化性能は、転移学習した EfficientNet の方が優れていると判断した。

Table 4 Evaluating of classification accuracy of transferred EfficientNet

class	accuracy	F_1 score	macro F_1
grade A	0.652	0.557	0.591
grade B	0.646	0.594	
grade C	0.871	0.623	

しかしながら、転移学習した EfficientNet の汎化性能においても、果実生産現場では通常利用できるほどのクラス分類精度には達していないと考える。今回、モデルの分類精度が低くなった原因として考えられるのは、次の 2 点である。

- (1) 果実 3 等級のラベル付けの際に、明確で客観的な分類基準を与えることができない。
- (2) 画像データセットサイズが小さいことにより、画像の特徴量抽出が上手くいかず、過学習も生じてしまった。

(1) に関しては、主観的クラスラベリング問題であり、客観的な分類基準の定義を行う、またはラベラーのラベル付けの誤差の問題として捉え、この誤差を最小化するようなアルゴリズムが必要と考える。(2) に関しては、共同研究企業の株式会社久留米原種育成会の協力を得て、モデル構築に用いる果実画像を増やしていくこととする。

5. 果長の推定

横並びに置かれた果実の果長の推定を試みた。果長の推定手順は次のとおりである。

- (1) 前章で利用した物体認識モデルにより、果実および園芸用ラベルの検知を行う。
- (2) 検知した果実および園芸用ラベルを含む矩形領域 (図 8(a)) を画像から切り出す。
- (3) 切り出した果実および園芸用ラベル画像の縦幅のピクセル数、横幅のピクセル数と園芸用ラベルの長さの実測値から果長を推定する。推定には次式を用いた。

$$L = \frac{px_c}{px_l} \times 8.8 \quad (4)$$

ここで、 L は果長の推定値 (cm)、 px_c は対象果実画像の縦幅のピクセル数、 px_l は園芸用ラベル画像の横幅のピクセル数、定数 8.8 は園芸用ラベルの長さの実測値 (8.8 cm) を表す。合計 50 本の果実を含む画像データを用いて、本推定モデルの果長推定を行った結果を図 8(b) に示す。グラフの横軸は果実番号、縦軸は果長の実測値と推定値の誤差である。実測値と推定値の誤差の二乗平均平方根 (RMS: Root Mean Square) を算出すると、約 0.56 cm となった。キュウリ果実の生産現場では、果長は cm 単位での計測が主であり、mm 単位での計測はほとんど行われない。従って、本推定モデルの推定精度は実用面で許容範囲内であると考えられる。

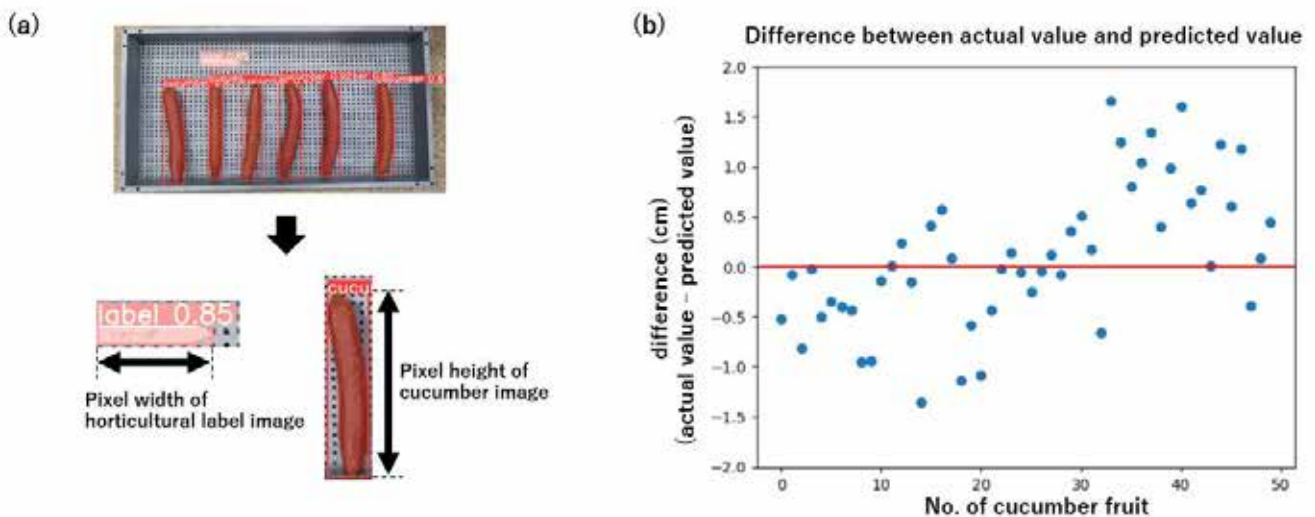


Fig. 8 Method and results for estimation of cucumber fruit length

6. 結 言

本研究では、キュウリ果実画像データを用いて、果実の品質等級分類モデルと果長の推定モデルを構築した。品質等級分類モデルでは、YOLOv8 を用いた物体認識モデルと CNN モデルによるクラス分類を試み、モデルの汎化性能を検証した。果長の推定モデルでは、果実と園芸用ラベルの画像サイズをもとに、実際の果長を推定し、その精度を検証した。これらの結果と得られた結論をまとめると、次のとおりである。

- (1) 物体認識モデルによる画像内の果実位置の検出と検出範囲の精度は高く、検出タスクは問題なく行えた。
- (2) 転移学習させた EfficientNet モデルの果実等級分類精度が最も高かった。しかしながら、実用面を考慮すると、本モデルの分類精度でも十分と言えず、改善が必要と考えられた。
- (3) 生産現場で利用可能なほど、高い精度で果長の推定が可能となった。

地域や生産現場、農業協同組合等によってキュウリ果実の品質分類基準は異なり、現在の品質分類基準には統一的な定義がない。そこで、本研究を発展させ、得られたモデルによる定義を客観性の高いキュウリ果実品質分類定義として提案していきたい。そのためには、より多くの画像データセットの用意に加え、複数の品質分類基準を用いたラベリングを行うことで、ラベリング誤差を最小化し、分類精度の向上を目指したい。

謝 辞

本研究では、株式会社久留米原種育成会の皆様にご提供とご助言を頂きました。ここに謝意を示します。

文 献

- (1) 総務省統計局 (2022) 「日本の統計 2022」, <https://www.stat.go.jp/data/nihon/index2.htm> (2023. 12.30 アクセス)
- (2) 農林水産省 「農業労働力に関する統計」, <https://www.maff.go.jp/j/tokei/index.html> (2023. 12.30 アクセス)
- (3) Liakos et al., “Machine learning in agriculture: A review”, *Sensors*, 18(8), 2674, 2018
- (4) Yoosefzadeh-Najafabadi et al., “Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean”, *Front. Plant Sci.*, 11:624273, 2021
- (5) Roboflow, Roboflow: Give your software the power to see objects in images and video, [online]<https://roboflow.com/> (2023. 12.30 アクセス)
- (6) Ultralytics, ホーム - Ultralytics YOLOv8 ドキュメント, [online]<https://docs.ultralytics.com/> (2023. 12.30 アクセス)
- (7) Simonyan & Zisserman, “Very deep convolutional networks for large-scale image recognition”, *arXiv:1409.1556*, 2014
- (8) Tan & Le, “EfficientNet: Rethinking model scaling for convolutional neural networks[C]”, *Proceedings of the 36th International Conference on Machine Learning*, PMLR 97, 2019