

転移学習によるアボカドの食べ頃分類手法と 応用システムの開発

杉本 隼斗 濱川 礼²

概要: 本論文では、転移学習によるアボカドの食べ頃分類手法とそれを応用した一般消費者向けの応用システムについて述べる。アボカドの食べ頃は果皮の色などをもとに一般消費者が判別する必要があるが、正確な見極めは熟練者でなければ難しい。本研究では、近年の深層学習技術の発展とスマートフォンなどの画像キャプチャデバイスの普及に着目し、転移学習による画像分類を用いたアボカドの食べ頃分類手法を提案する。アボカドの写っている画像を入力とし、未熟、食べ頃、過熟の3クラスに識別するように独自のデータセットの構築と分類器の学習を行う。応用システムは一般消費者が手軽に扱えるアプリケーションであることを重視しスマートフォンアプリの開発を行う。評価実験では、事前学習済みモデルから得られる特徴量を利用したニューラルネットワークによる分類が作成した2つのデータセットにおいて最も高い精度を示し、提案手法はデータセットに依存しない有用性を示した。開発した応用システムも被験者から高い評価を得ることができ、システムの有用性が明らかになった。

キーワード: アボカド, 転移学習, 画像分類

Avocado Ripeness Classification System Using Transfer Learning

HAYATO SUGIMOTO REI HAMAKAWA²

Abstract: In this paper, we describe a method for classifying avocado based on transfer learning and an application system for consumers. It is necessary for consumers to determine the eating time of avocados based on the color of the skin and other factors, but it is difficult for only skilled consumers to accurately determine the eating time. In this study, we propose a method for classifying avocados based on the recent development of deep learning technology and the widespread use of image capturing devices such as smartphones. Using an image of an avocado as an input, we construct a unique dataset and train a classifier to discriminate between three classes: unripe, ripe, and overripe. The application system is developed as a smartphone application, with an emphasis on being an application that can be easily handled by consumers. In the evaluation experiments, the classification by the neural network using the features obtained from the pre-trained model showed the highest accuracy in the two datasets created, indicating the usefulness of the proposed method independent of the dataset. The developed application system was also highly evaluated by the test subjects, demonstrating the usefulness of the system.

Keywords: Avocado, Transfer learning, Image Classification

1. はじめに

アボカドは栄養価が高く、世界的に需要が拡大している

南方果実である。その栄養価の高さとカロリーの低さからスーパーフードとして知られている。世界のアボカド市場において中心的品種であるハスアボカドは収穫後から食べ頃になるにつれて果皮の色が緑色から茶色、黒紫色に変化する [1]。スーパーマーケットなどの小売店では未熟の緑色の状態のアボカドも販売されているため、一般消費者が購

¹ 中京大学大学院 工学研究科情報 工学専攻
Graduate School of Engineering, Chukyo Unibersity

² 中京大学 工学部 情報工学科
School of Engineering, Chukyo Unibersity

入時や調理時にアボカドの食べ頃を判別する必要がある。しかし、この判断指標は曖昧であり正確な判別には多くの経験が必要とされる。ファーマインド社によって2014年に行われたアボカドのイメージに関するアンケート調査 [2] によると、67%の回答者が「食べ頃がわかりにくい」と答えている。食べ頃の判断を誤ってしまった場合、食べ頃を迎える前の未熟な状態では実が硬く、食べ頃を過ぎて追熟を続けた過熟な状態では実に黒ずみなどの異常が発生し、食品ロスに繋がる恐れがあるため、このような調査結果は深刻な問題である。一部の小売店では食べ頃の判断を支援するために、アボカドの追熟過程がわかるカラーチャートを店頭で設置する工夫がされている。しかし、アボカドは個体や保管方法によって果皮の色や質感の変化が異なるため、カラーチャートのみを参考にするだけでは食べ頃の見極めに失敗する恐れがある。握ったときの硬さも加味することで判別の正確性は向上するが、店頭で多くの人々が握って確かめることで商品の劣化に繋がることから推奨されるべき行為ではない。このような問題がある中で、アボカドの食べ頃を分類する研究がされているが、これらの研究によって提案されている手法 [3], [4] は一般消費者を利用対象としておらず、食べ頃を知るために特殊な機器を利用する必要がある。一般消費者が容易に利用可能な手法は少なく、一般消費者が食べ頃を見極めるのが難しいという問題の解決には至っていない。そこで本研究室では画像認識をベースとした一般消費者向けのアボカドの食べ頃分類システムの開発を行ってきた [5], [6]。画像認識をベースとした手法はスマートフォンなどの普及率の高いデバイスから利用可能なため、手軽さの点で非常に有用である。[7] では同様に画像認識をベースとした手法が使われているが色成分から得られる特徴量とテクスチャ特徴量を分類に使用する従来の手法により食べ頃分類を行なっているため、深層学習が発展した現代において分類精度改善の見込みがあると考えられる。本研究室で行なった研究の [5] はデータレベルでの改善が必要だったのに対し、[6] は深層学習を活用した手法により成果を得たが、本論文では転移学習を利用することで、小規模データセットに対してより安定した分類精度の手法とモデルの開発を試みる。深層学習では、一般的に大量のデータが必要とされるが、ImageNet などの大規模データセットの事前学習済みモデルを活用する転移学習手法は、比較的小規模なデータセットにおいても画像認識精度向上の効果があることが報告されている。ウメ、リンゴ、オレンジなどの果物を対象にした食べ頃分類に関する研究 [8], [9] においても転移学習による手法が利用されている。アボカドを対象とした大規模な公開データセットは存在しないため、本研究においても小規模なデータセットに対応可能な転移学習による手法が有用だと考えた。

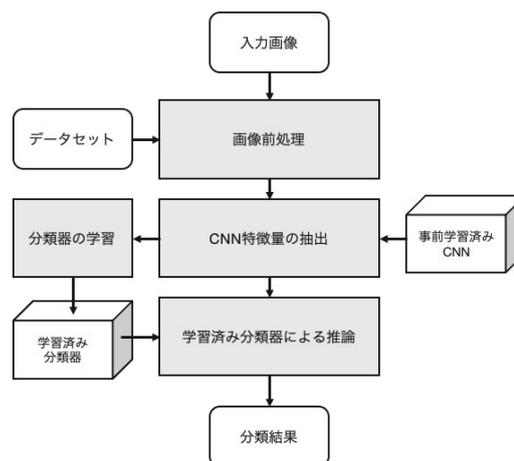


図 1 手法の流れ。

Fig. 1 Method overview.

2. 提案手法

本論文では、転移学習による画像分類を用いたアボカドの食べ頃分類手法を提案する。図1に本手法の流れを示す。画像入力から画像前処理を行い、次に事前学習済み CNN から特徴量の抽出する。本論文では、ここで得られる特徴量を「CNN 特徴量」と呼ぶ。その特徴量を未熟、食べ頃、過熟の3クラスに分類するように分類器の学習を行い推論に使用する。

2.1 転移学習

転移学習は事前学習済みモデルを活用し新たなタスクに適用する学習方法のことを指す。一般的に深層学習を用いた手法は大量のデータが必要とされるが、転移学習は大規模なデータセットで学習されたパラメータを使用して追加学習するため、小規模なデータセットにおいても高い精度でタスクを解決することができる。近年では ImageNet や MS COCO など大規模な公開データセットが存在するが、現実の多くのタスクでは公開データセットは存在せず、大量のデータを確保することは難しい。そのため転移学習のような小規模なデータセットに有効な手法は多くの研究で活用されている。

画像認識分野では、一般物体認識用大規模データセットである ImageNet の 1000 クラス分類を学習した CNN の事前学習済みモデルが TensorFlow や PyTorch などの主要な深層学習ライブラリで公開されている。これらのモデルに画像を入力して得られる特徴量は、画像の一般的な特徴を捉えており汎用性が高く、その特徴量を使用することで様々なタスクの分類器を学習することが可能である。

大量のアボカドの画像と食べ頃に関するラベルを取得することは困難であるため、転移学習のように小規模なデータセットであっても高い分類精度が期待できる手法が本研究に適していると考えられる。本研究では、ImageNet の

1000 クラス分類を事前学習済みの VGG16[10] を使用して CNN 特徴量を抽出する。VGG16 は 2014 年の VGG という ILSVRC においてローカリゼーションと分類タスクにおいて高い精度を達成した CNN のひとつである。VGG は複数個のスタックされた畳み込み層のブロックと分類器の役割を持つ 3 層の全結合層で構成される。シンプルにスタックされたネットワークのため転移学習などで利用がしやすいネットワークである。

2.2 画像前処理

画像前処理では、画像のリサイズとスケーリングを行う。事前学習済み CNN を使用する転移学習を行う際、CNN 特徴量の抽出に使用するネットワークの入力サイズに合わせた画像のリサイズとスケーリングを行う必要がある。VGG16 のネットワークの入力サイズは 224×224 の画像のため、センタークロップを行った後に 224×224 の大きさに画像をリサイズする。ここでセンタークロップを行う理由は、単純に画像をリサイズするとアボカドの形が変形してしまうだけでなく質感の情報を失う恐れがあり、それを回避するためである。スケーリングは VGG16 の事前学習時の処理と同様にデータセットの平均画素値を各画素から減算する処理を行う ($R = 123, G = 114, B = 104$)。

2.3 CNN 特徴量の抽出

CNN 特徴量を抽出するには、前処理された画像を事前学習済みモデルに入力しネットワークの中間層の出力を取得する。本研究では、図 2 のように VGG16 から CNN 特徴量を抽出する。ここで、ネットワーク中のどの層の出力を特徴量として使用するかによって特徴量の表現する情報が異なり分類性能に影響する。CNN が入力に近い層ではエッジやテクスチャなど一般的な画像表現を学習し、出力に近い層ではオブジェクトなどデータセットのドメイン固有の情報を学習することが明らかになっている [11]。そのため、事前学習済みモデルが学習したデータセットにドメインが類似していない場合は、分類器だけでなく一部の畳み込み層の再学習を行うこともある。これは転移学習の中でも fine-tuning と呼ばれている [12]。しかし、学習するパラメータ数が増加することから必要とされるデータ数は多くなってしまったため、本研究では畳み込み層の再学習はせずに VGG16 の 2 番目の全結合層から得られる出力を CNN 特徴量として抽出する。抽出した特徴量の大きさは 4096 次元になる。

2.4 分類器の学習と推論

分類器によって推論結果を得るために、分類器は CNN 特徴量を未熟 (unripe)、食べ頃 (ripe)、過熟 (overripe) の 3 クラスに分類するように学習させる。推論時、学習した分類器は各クラスに該当する確率を出力し、これを食べ頃

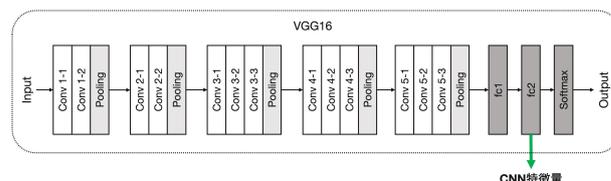


図 2 VGG16 から CNN 特徴量を抽出する方法。

Fig. 2 A method for extracting CNN features from VGG16.

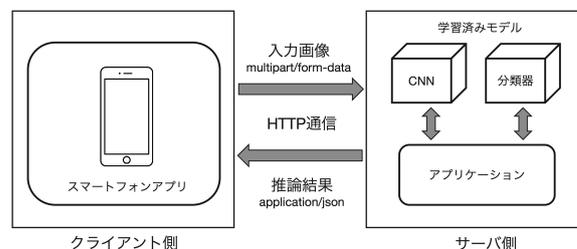


図 3 システム構成。

Fig. 3 System architecture.

の分類結果として使用する。分類器はガウスクernelによる SVM, ランダムフォレスト, ニューラルネットワークを用いて学習を行った後に交差検証を実施し最も精度の良かったものを選択する。

3. 応用システム

本研究では一般消費者がより手軽に利用可能なアプリケーションを目指し、スマートフォンアプリを中心とした応用システムの開発を行なった。[6]で開発したアプリケーションとシステムの構成が異なる。

3.1 構成と実装

図 3 にシステム構成図を示す。本システムはユーザが操作するスマートフォンアプリと学習モデルを用いて推論を行うサーバアプリケーションから構成されている。

ユーザによって本スマートフォンアプリ上で撮影された画像または端末内から選択された保存済み画像は、HTTP 通信の POST メソッドを使用し、multipart/form-data 形式の画像データとしてサーバアプリケーションに送信される。サーバアプリケーションは送信された画像データをもとに、2 章で述べた画像前処理、CNN 特徴量の抽出、分類器による推論の処理を行い、分類結果を JSON 形式でレスポンスする。最後にレスポンスされた分類結果をもとにスマートフォンアプリで食べ頃を可視化表示する。分類器は学習実験で最も性能の高かったモデルを採用している。

スマートフォンアプリはクロスプラットフォームなアプリ開発が可能なフレームワーク Flutter、サーバアプリケーションは Python の Web 開発フレームワーク Flask を使用して開発した。画像前処理には OpenCV、CNN 特徴量の抽出に使用する事前学習済みモデルは TensorFlow で公開されているモデルを使用している。

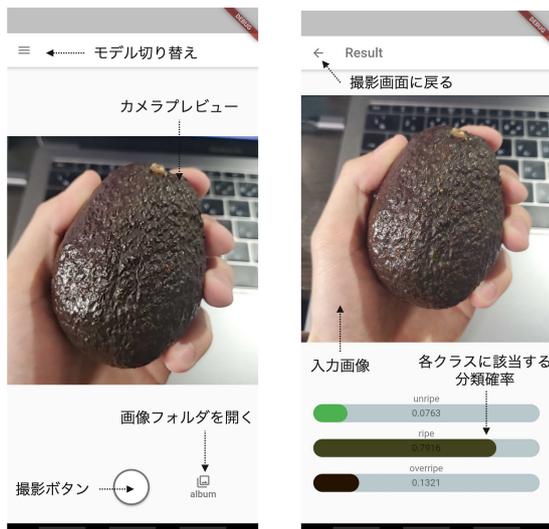


図 4 撮影画面 (左). 結果表示画面 (右).

Fig. 4 Camera preview screen (left). Result screen (right).

3.2 画面構成

スマートフォンアプリは撮影画面と結果表示画面の2画面で構成されている。それぞれの実際の画面を図4に示す。撮影画面では、食べ頃を分類したいアボカドの画像を撮影もしくは保存された画像から選択することができる。ユーザによって画像が決定された後に自動的にサーバアプリケーションにデータが送信され、結果表示画面へと遷移する。結果表示画面では、分類に使用された入力画像と分類結果を示すパーセンテージバーが表示される。パーセンテージバーを確認することで分類器が示す各クラスに該当する確率がわかる。これにより一般消費者はアボカドの食べ頃に関する情報を得ることができる。

4. 評価実験

本研究では、転移学習手法の有用性を検証するための交差検証による分類精度評価、応用システムの有用性を検証するための被験者評価実験の2つの実験を行なった。

4.1 データセット

本研究では2種類のデータセットを構築した。1つ目のデータセットは画像キャプチャデバイス、照明などの撮影環境を固定して撮影した画像をもとに構成される固定環境データセット、2つ目のデータセットは複数のスマートフォンで撮影された画像をもとに構成される非固定環境データセットである。クラスは未熟、食べ頃、過熟の3クラスである。表1に実際に収集したデータ数を示す。

表 1 データセットの内訳。

Table 1 The number of data in each dataset.

データセット	未熟	食べ頃	過熟	サンプル数
固定環境	1582 枚	506 枚	452 枚	223 個
非固定環境	87 枚	128 枚	82 枚	62 個

4.1.1 固定環境データセット

固定環境データセットは、アボカドの食べ頃分類を行う際に撮影環境が影響し分類精度が安定しない可能性を考慮し、それらの原因を排除して分類器の学習と手法の比較評価を行うことを目的に構築した。本研究の実験で用いた固定環境データセットは[6]で構築したデータセットに新たにデータを追加したものである。

4.1.2 非固定環境データセット

非固定環境データセットはシステム運用時により近いデータを収集することを目的として構築した。ユーザはスマートフォンアプリで画像を撮影することが想定されるため、スマートフォンで撮影した画像から構成されている。撮影に使用したスマートフォンは Galaxy S10, Google Pixel 4a, iPhone SE2 の3種類である。撮影時の照明や背景などの外部環境は、固定環境データセットのように統一はせず、室内の様々な環境において撮影した。ラベル付与は破壊検査が困難であったため、固定環境データセットとは異なり果皮の目視評価のみで行なった。

4.2 分類精度評価

本節では、構築したデータセットを使用し学習した分類器の性能評価を行なった評価実験について述べる。各実験で分類器の正解率を算出し、その結果に対する考察を行なった。

4.2.1 実験方法

分類精度評価では転移学習による本手法と色成分とテクスチャ特徴をもとに分類を行う従来手法と比較実験を行った。比較手法では、 $L^*a^*b^*$ 色空間の各色成分から作成した色ヒストグラムと GLCM 共起行列によって得られるテクスチャ情報を持つ Haralick 特徴量を取得し、それらを結合したものを分類器に学習させる。

各モデルの性能評価はグループ k 分割交差検証より行なった。グループ k 分割交差検証は、 k 分割交差検証においてデータセットを k 個のセットに分割する前にデータをグループ化し、同じグループのデータは必ず同じセットに割り振られるという制約のもと交差検証を行う方法である。本研究で構築したデータセットは同一のアボカドから複数の画像を撮影しデータを収集しているため、通常の k 分割交差検証によってランダムにデータセットを分割した場合、同じアボカドから撮影した画像で同じラベルを持つデータが訓練用セットと検証用セットの両方に割り振られる恐れがある。画像の撮影日が異なれば追熟によってラベルが一意に定まることはないため問題ないが、そうでない場合、検証時に既知のデータに酷似したデータを使用することになり正確に分類器の汎化性能を評価することができない。一方でグループ k 分割交差検証は検証データに未知のデータが割り振られるため正確に汎化性能を評価することができる。本研究では、同じアボカドから撮影した画像

表 2 交差検証の結果: 固定環境データセット.

Table 2 The result of cross-validation: Restricted Dataset.

特徴量	分類器	正解率
色成分+テクスチャ	SVM	83.6
	ランダムフォレスト	84.2
	ニューラルネットワーク	85.0
CNN 特徴量	SVM	83.7
	ランダムフォレスト	84.3
	ニューラルネットワーク	85.8

表 3 交差検証の結果: 非固定環境データセット

Table 3 The result of cross-validation: Non-Restricted Dataset.

特徴量	分類器	正解率
色成分+テクスチャ	SVM	77.2
	ランダムフォレスト	79.0
	ニューラルネットワーク	80.7
CNN 特徴量	SVM	82.4
	ランダムフォレスト	88.4
	ニューラルネットワーク	89.4

でも追熟の経過次第では異なるラベルを持つ場合があるため、個体と撮影日の2つの基準をもとにグループ化し交差検証を行った。さらにグループ化後、各クラスのグループ数を揃えるダウンサンプリングを実施しデータセットの不均衡を解消した。

4.2.2 結果と考察

各データセットを使用したグループ5分割交差検証の結果を表2と表3に示す。固定環境データセットを用いた実験では、CNN 特徴量を利用したニューラルネットワークによる分類が正解率 85.8%と最も高く、非固定環境データセットを用いた実験でも同様に CNN 特徴量を利用したニューラルネットワークによる分類が正解率 89.4%と最も高かった。このことから VGG16 の中間層の出力から得た CNN 特徴量を用いた転移学習がアボカドの食べ頃分類において有効な手法であったと言える。

使用した特徴量で正解率を比較したとき、固定環境データセットは転移学習による正解率の向上がわずかに 0.8%だったが、非固定環境データセットでは 8.7%と大きな向上が見られた。また、データセット間で精度の変化を見たとき、色成分から得られる特徴量とテクスチャ特徴量を用いた分類器の学習では固定環境データセットを使用したときと非固定環境データセットを使用したときで比べると正解率が低下しているのに対して、CNN 特徴量を利用した転移学習では正解率が向上している。CNN 特徴量が撮影環境などに影響せず汎用的な画像特徴を取得できているため、データセットに左右されず分類精度が安定していると考察できる。CNN 特徴量を利用した転移学習は特徴量抽出の段階でデータセットに特化した処理が不要であり簡単に利用することができるため実装面から見ても有用な手法と考えら

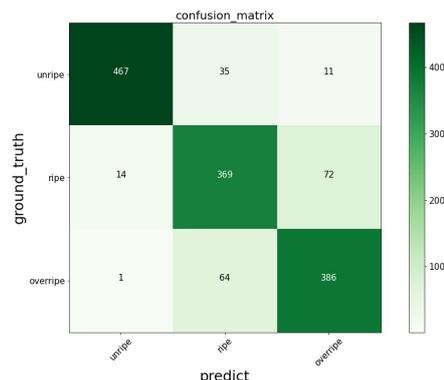


図 5 混同行列: 固定環境データセット.

Fig. 5 Confusion matrix: Restricted Dataset.

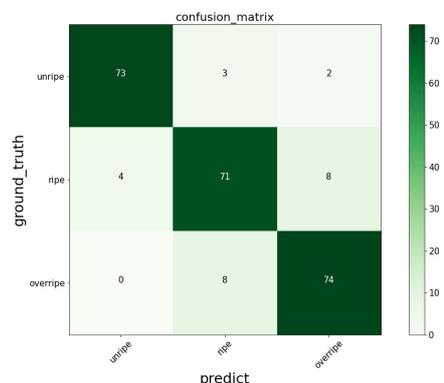


図 6 混同行列: 非固定環境データセット.

Fig. 6 Confusion matrix: Non-Restricted Dataset.

れる。図5と図6は最も正解率の高かった CNN 特徴量とニューラルネットワークの組み合わせで再度交差検証を一度行い、その分類結果を混同行列にしたものである。どちらのデータセットにおいても食べ頃クラスと過熟クラスの間での誤分類が多く発生していた。これは未熟から食べ頃になるときに比べ、食べ頃から過熟になるときに果皮の色の変化が少ないことと過熟クラスのデータ数が少ないことが影響していると考えられる。

4.3 システムの有用性評価

本節では、開発した応用システムの有用性を明らかにするために行った被験者実験について述べる。

4.3.1 実験方法

大学生 11 人に対して、カラーチャートと本システムを利用してアボカドの食べ頃判別をしてもらい、その後アンケート調査を行なった。6 個の 5 段階評価設問、自由記述式設問を用意した。画像はスマートフォンアプリ内で撮影してもらったものを使用する。

4.3.2 結果と考察

アンケート調査の結果を表4に示す。分類精度に関する Q5 以外の項目で 4 点以上の得点を得た。実際にシステム

表 4 アンケート調査の結果

Table 4 Results of the questionnaire survey.

番号	質問	1	2	3	4	5	平均
Q1	食べ頃を判断する際に役に立つか (主観的に需要を感じるか)	0	0	2	5	4	4.2
Q2	食べ頃を判断する際に役に立つか (客観的に需要を感じるか)	0	0	0	4	7	4.6
Q3	可視化される情報は十分だと思ったか	0	1	1	4	5	4.2
Q4	システムを手軽に利用できると感じか	0	0	0	1	10	4.9
Q5	本システムは食べ頃判断支援を目的としたときに 分類精度は十分だと感じたか	0	0	4	4	3	3.9
Q6	カラーチャートによる支援に比べて良いと思ったか	0	0	2	6	3	4.1

を運用すると撮影時の環境によって精度が安定せず、アボカドの外観や硬さとは反する分類結果が出ることもあり、今後最も改善すべき点として挙げられる。システムの手軽さに関する Q4 では 11 人中 10 人の被験者から 5 点という最高点の回答を得た。さらに「システムの良かった点」という自由記述設問において「カメラで撮影するのみで使用可能な点」、「操作が簡単な点」など同様の意見を得ることができた。カラーチャートと比較した Q6 においても平均が 4.1 点と高い点数を得ることができ、本システムの有用性を示した。しかし、「システムの改善点」という自由記述設問において「はっきりと食べ頃かそうでないのかわからない」、「2 つのパーセンテージバーが伸びたときどうすれば良いかわからない」という意見の回答をもらい、課題となっているカラーチャートの曖昧性の問題は完全に解消することはできなかった。

5. おわりに

アボカドの食べ頃判別支援を目的として転移学習による食べ頃分類手法の提案と応用システムの開発を行なった。転移学習による分類手法はデータセットに依存せず汎用的な性能を示した。提案手法を実装した応用システムではスマートフォンアプリとサーバアプリケーションの開発を行い、ユーザが手軽にアボカドの食べ頃判別支援を受けることができるようになった。被験者による評価実験では、高い評価を得ることができシステムの有用性が明らかになった。しかし交差検証で得られた分類精度と実際の運用時の分類精度には差があったため、今後はさらにデータ収集を行い安定した分類器の学習と厳密な評価を行う必要がある。小売店の売り場などでのシステム利用を考えたとき、複数個のアボカドから食べ頃のアボカドを見分けたいことがある。今回の提案手法は画像分類を用いた手法であることから画像内の一つのアボカドの食べ頃に関する情報しか得ることができないため、今後は物体検出などを適用し複数個のアボカドの食べ頃を見分ける技術を開発したい。

参考文献

- [1] ACox, K., McGhie, T. K., White, A. and B Woolf, A.: Skin colour and pigment changes during ripening of 'Hass' avocado fruit, *Postharvest Biology and Technology*, Vol. 31, No. 3, pp. 287-294 (2004).
- [2] 株式会社ファーマインド: アボカド食べ頃パンフレット. https://www.farmind.co.jp/corporate/business/catalog/avocado_catarog.pdf (参照 2022-2-16).
- [3] Monzurul Islam, K. W. and Dinh, A.: Assessment of Ripening Degree of Avocado by Electrical Impedance Spectroscopy and Support Vector Machine, *Journal of Food Quality* (2018).
- [4] Pinto, J. E., Rueda-Chacon, H. and Arguello, H.: Classification of hass avocado (*persea americana* mill) in terms of its ripening via hyperspectral images, *TecnoLógicas*, Vol. 22, No. 45, pp. 109-128 (2019).
- [5] 杉本隼斗, 久野文菜, 谷口航平, 濱川礼: 深層学習を用いたアボカドの追熟段階分類手法の提案, 人工知能学会全国大会 (2020).
- [6] 杉本隼斗, 濱川礼: エンドユーザ向けアボカド食べ頃分類モバイルアプリにおける深層距離学習を利用した分類手法の検討, 研究報告ヒューマンコンピュータインタラクション, No. 5 (2021).
- [7] Jaramillo-Acevedo, C. A., Choque-Valderrama, W. E., Guerrero-Álvarez, G. E. and Meneses-Escobar, C. A.: Hass avocado ripeness classification by mobile devices using digital image processing and ANN methods, *International Journal of Food Engineering*, Vol. 16, No. 12 (2020).
- [8] 建本聡, 原田陽子, 今井健司: 深層学習を利用したウメ「露茜」の画像による熟度分類, *農業情報研究*, Vol. 28, No. 3, pp. 108-114 (2019).
- [9] Hermana, A. N., Rosmala, D. and Husada, M. G.: Transfer Learning for Classification of Fruit Ripeness Using VGG16, *The 4th International Conference on Computers in Management and Business*, p. 139-146 (2021).
- [10] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, *International Conference on Learning Representations* (2015).
- [11] Zeiler, M. D. and Fergus, R.: Visualizing and Understanding Convolutional Networks, *Computer Vision – ECCV 2014*, Springer International Publishing, pp. 818-833 (2014).
- [12] 中山英樹: 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*, Vol. 115, No. 146, pp. 55-59 (2015).