

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

杉本 隼斗*¹ 濱川 礼*²

Investigation of deep metric learning for mobile application that classifies avocado ripeness for end users.

Hayato Sugimoto*¹, Rei Hamakawa*²

Abstract – It is said that the eating time of an avocado can be determined using the color, texture, and firmness of the rind as indicators, but it is difficult to determine the eating time of an avocado with high accuracy only for experienced users, which is a problem faced by end users of avocados. In this study, we developed a deep learning model that classifies avocados into three classes (unripe, ripe, and overripe) based on image input, and implemented a mobile application that can be run on a smartphone equipped with the model. For the deep learning model, we investigated a classification method using deep metric learning. Deep metric learning has achieved many successes in face recognition tasks. It is considered useful when applying deep learning to datasets with small differences in image features between classes and a small amount of data for each class, and the dataset we collected in this study has the same characteristics. The model was able to classify the evaluation data with an accuracy of 89.77%.

Keywords : avocado, deep metric learning, mobile application

1. はじめに

アボカドは栄養価の高い追熟型の果物である。追熟型の果物は未熟の状態では収穫され、適切な追熟を経て食べ頃を迎える。アボカド市場において中心的な品種であるハスアボカドは追熟によって果皮の色が緑色から茶色、黒紫色に変化していき食べ頃を迎えるという特徴がある^[1]。このような外観の変化をエンドユーザが判別し、食べ頃を見極める必要があるが、指標が曖昧なことや個体差があることから、熟練者でなければ正確な見極めが難しいことが課題である。そこで、本研究ではこのようなエンドユーザが抱える課題を解決するため、スマートフォンやタブレットなど汎用デバイスで実行可能なアボカドの食べ頃を分類するモバイルアプリの開発を行った。

アボカドの食べ頃を分類する既存研究では、特殊なデバイスを必要とする手法が多く提案されている^{[2],[3]}。そのような手法を汎用デバイス上に実装することは困難なことから、エンドユーザが食べ頃を見極めるのが難しいという課題の解決には至っていないのが現状である。近年では、機械学習を用いた画像認識手法によって硬度を予測する研究^[4]がされており、スマートフォンやタブレットなど汎用デバイスに搭載されたカメラを使用することで実行可能なためエンドユーザを対象

としたシステムにおいても有効だと考えられる。そこで、本研究でも同様に画像認識による手法を用いることとした。アボカド以外の果物では、CNNによる深層学習の画像認識技術を用いた手法が食べ頃分類や収穫時期の判別において多くの成果を上げている^{[5]~[8]}。CNNを用いた手法は多様な課題に対し、高精度かつ頑強性の高い分類が可能とされているため、本研究ではCNNをベースとしたアボカドの食べ頃分類手法を用いる。しかし、CNNなど深層学習を利用した手法は大量のデータがなければ学習を行うのが困難という問題がある。アボカドの食べ頃に関する教師データを集めるのは高コストなため、大量のデータを収集することは現実的ではない。そこで、顔認識タスクなどで成果を上げている深層距離学習^[9]に着目した。深層距離学習は分類対象のクラス間における画像特長の差が小さく、各クラスのデータ数が少量のデータセットを用いた学習など難しい問題設定の画像認識に対して有効とされている^[10]。これらの特徴は本研究で扱うデータセットにおいても共通していると考え、深層距離学習を利用した分類手法の検討を行った。

2. アボカドの食べ頃分類手法

図1に手法全体の流れについて示す。画像入力からCNNによって3クラス(未熟, 食べ頃, 過熟)に結果を分類する。構築したCNNはResNet-18^[11]をベースにしたネットワークで3クラスに分類結果を出力できるようにSoftmaxを活性化関数に持つノード数3の

*1: 中京大学大学院 工学研究科 情報工学専攻

*2: 中京大学 工学部 情報工学科

*1: Graduate School of Engineering, Chukyo University

*2: School of Engineering, Chukyo University

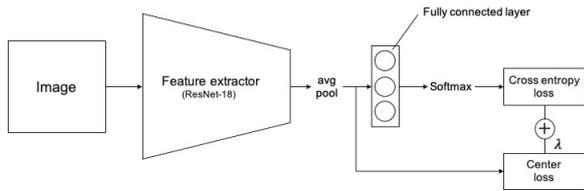


図1 手法の流れ
Fig. 1 Overview of the method.

出力層を新たに作成し追加した。学習時の損失関数に深層距離学習の一手法とされる Center loss^[12]を導入し、深層距離学習の効果を検証した。学習に使用するデータセットは独自に作成を行った。

2.1 深層距離学習

本研究で使用する深層距離学習の手法は Center loss と呼ばれるものである。Center loss では、一般的な分類タスクにおける CNN の損失関数に、特徴空間を最適化する項を追加する。CNN の特徴抽出器から得られる特徴量ベクトルをもとに、特徴空間における各データ点と各クラスの中心点の距離の最適化を行う。特徴空間における各データ点を x_i 、各クラスの中心点を c_{yi} 、ミニバッチサイズを m としたとき、Center loss の項 \mathcal{L}_c は (1) 式で表される。

$$\mathcal{L}_c = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{yi}\|_2^2 \quad (1)$$

最終的な損失関数 \mathcal{L} は交差エントロピーと softmax による分類誤差を \mathcal{L}_s としたとき、ハイパーパラメータ λ を用いて (2) 式で表される。

$$\mathcal{L} = \mathcal{L}_s + \lambda \mathcal{L}_c \quad (2)$$

j 番目のクラスの中心点 c_j はハイパーパラメータ α を用いて (3)(4) 式によって更新される。

$$\Delta c_j = \frac{\sum_{i=1}^m \delta(y_i = j) \cdot (c_j - x_i)}{1 + \sum_{i=1}^m \delta(y_i = j)} \quad (3)$$

$$c_j^{t+1} = c_j^t - \alpha \cdot \Delta c_j^t \quad (4)$$

ハイパーパラメータ λ と α は実験的に適した値を見つける必要がある。それらの値を変化させてモデルを学習し、本研究における深層距離学習の有効性を検証する。

3. 実験と評価

エンドユーザに対してアボカドの食べ頃見極めを支援するには、高精度な分類を可能とする手法が必要とされる。そこで本手法で作成したモデルの性能調査を行った。さらに、深層距離学習を使用しなかった場合

における性能と比較を行い、本研究の分類手法に深層距離学習を導入することが有効かどうかを検討した。

3.1 データの収集

画像の撮影には、Aoboco 製の 200 万画素で撮影可能なカメラと固定用の三脚、SAMTIAN 製の小物撮影ボックスを使用した。その様子を図 2 に示す。小物撮影ボックスは照明条件など外部からの影響を受けないよう撮影環境を固定するために使用した。ラベル付与は外観の目視評価と食味評価をもとに行った。各クラスにおけるデータの一例を図 3 に示す。訓練に使用する画像と評価に使用する画像は異なるアボカドから撮影した。クラスごとの画像枚数とアボカドのサンプル数を表 1 に示す。

3.2 学習条件

本節では学習時のハイパーパラメータやデータ拡張の条件について記述する。ネットワークの入力サイズは 320×320 、バッチサイズは 16 に設定し、最適化手



図2 データ収集時の画像撮影環境
Fig.2 Imaging environment for data collection.

表1 収集したアボカドの画像とサンプルの数
Table 1 Collected images and samples of avocado.

	未熟	食べ頃	過熟	サンプル
訓練用	1458 枚	436 枚	367 枚	187 個
評価用	70 枚	70 枚	85 枚	34 個



図3 各クラスの画像例 (左から未熟, 食べ頃, 過熟)
Fig.3 Sample images of each classes. (from left to right: unripe, ripe, overripe)

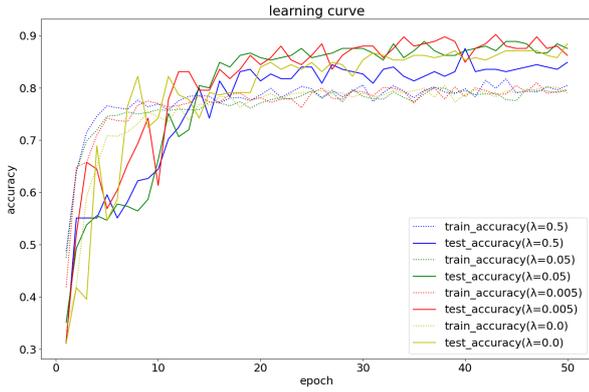


図4 λ を変化させたときの Accuracy のグラフ ($\alpha=0.2$)

Fig. 4 Accuracy curve.

法は学習率を 10^{-6} の RMSprop を用いた。これらの値は事前実験をもとに決定した。Center loss に関するハイパーパラメータ λ と α は値を実験ごとに変化させ、モデルの性能の変化を観察した。データ拡張は左右反転、回転、拡大・縮小、せん断、水平・垂直シフトをミニバッチ生成時に各画像に対してランダムに施した。また、全ての画像をネットワークの入力サイズに合わせてリサイズし、正規化処理を行った。不均衡データ対策としてアンダーサンプリングを行った。最も画像数が少なかった過熟クラスの 367 枚に合わせて、他のクラスも学習に使用される画像はランダムにサンプリングされた 367 枚となる。

3.3 結果

λ を変化させモデルを学習した際の結果を図 4 に示す。値を変化させても Accuracy に大きな差が観察されることはなく、Center loss の有効性を示すことはできなかった。最も分類精度が高いモデルは $\lambda = A$, $\alpha = 0.2$ としたときに評価データに対する Accuracy は 89.77% に到達した。 α は値を変化させて実験を行ったところ、0.2 の値が適していることがわかった。

3.4 考察

Center loss を用いた深層距離学習の手法が本研究に対して有効でなかったのは、本研究の問題設定の易しさが原因と考えられる。深層学習は大量のデータが必要な手法であるが、必要なデータの数は問題設定の難易度によって異なる。本研究ではデータセット構築の際、画像収集を固定した環境で行ったため、深層距離学習を用いるには比較的分類が容易な問題設定であった可能性がある。そのため、各クラスが類似した画像であるものの、必要なデータ数は十分満たしており、通常分類損失のみを損失関数に用いた手法と大きな差が見られなかったと推測する。そこで、データセットの規模をさらに縮小し追加実験を行った。各クラスの訓練データ数を 100 枚の画像にしたときの結果を図

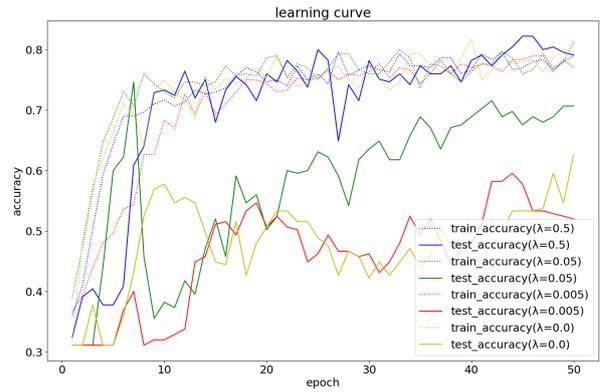


図5 縮小したデータセットにおいて λ を変化させたときの Accuracy のグラフ ($\alpha=0.2$)

Fig. 5 Accuracy curve for reduced dataset.

5 に示す。データ規模を縮小したとき、 $\lambda = 0.0$ のモデルは訓練データに対して過学習しており、 $\lambda = 0.5$ のモデルが他のモデルに比べ良い汎化性能を示していることがわかる。この結果から Center loss を用いた損失関数が、アボカドの食べ頃を分類する問題設定において、モデルが汎化性能を獲得するために必要なデータ数を確保できていない場合に有効だといえる。ここから、通常分類誤差のみでは過学習が起きてしまう場合に、Center loss は正則化手法の一つとしてモデルの過学習を抑える効果を持つと考察できる。

本研究で収集したデータは撮影環境を固定しているが、今後は様々な環境で撮影されたデータに対して実験を行いたいと考えている。環境を固定しなかったとき、問題設定の難易度が上がるため必要なデータ数が増え、過学習が起きる可能性が高くなると予想できる。その際、今回の実験で得られた結果をもとに、Center loss などの深層距離学習の手法を再度検討していきたい。

4. モバイルアプリ

クロスプラットフォーム実装が可能な Flutter を用いてモバイルアプリの開発を行った。図 6 に開発したモバイルアプリの流れを示す。ユーザはカメラで撮影

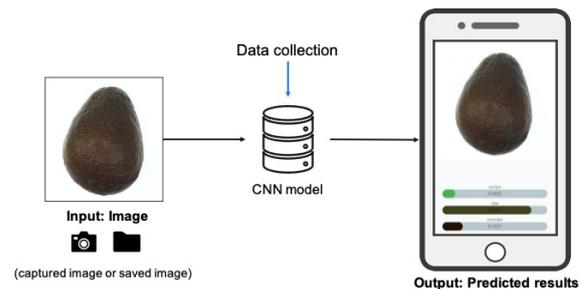


図6 モバイルアプリの流れ

Fig. 6 Overview of the mobile application.

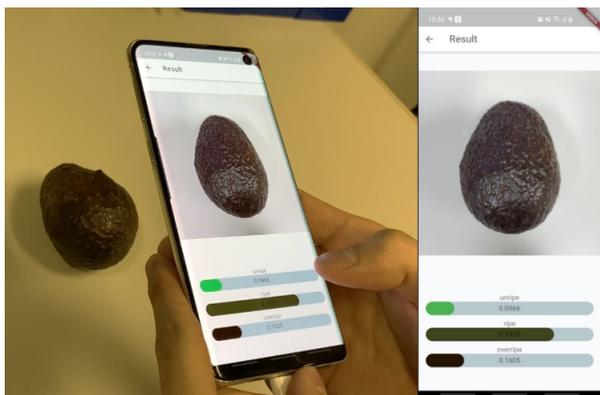


図7 使用の様子
 Fig. 7 The image when using mobile application.

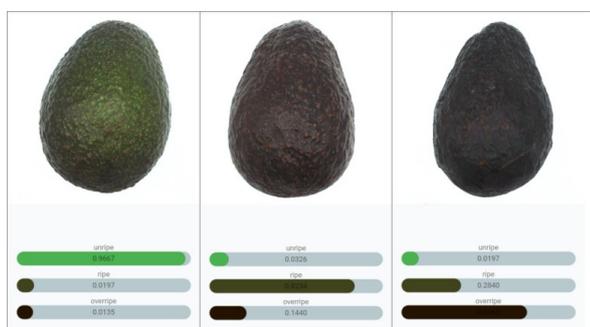


図8 食べ頃の分類結果
 Fig. 8 Ripeness Classification results.

した画像または事前に撮影された端末内に保存されている画像を入力することで3クラスの食べ頃の分類結果を得ることができる。分類結果は各クラスごとに用意されたプログレスバーによって可視化される。図7に使用の様子と図8に分類結果を示す。入力された画像に写るアボカドは外観から判断すると、もっともらしい分類結果であることが予想でき、実際に食味評価を行ったところモデルの分類結果と一致していた。

5. まとめ・今後の展望

本研究ではエンドユーザがアボカドの食べ頃を見極める際の支援を目的とした、食べ頃分類モバイルアプリの開発と、深層距離学習を用いた分類手法の検討を行った。実験によって最も性能の良かったモデルは89.77%のAccuracyを得た。さらに、深層距離学習の検討を行った結果、過学習が起きてしまうデータセット規模の場合に有効性を確認することができた。一方、本研究は撮影環境を再現した画像でなければ、高精度な分類を見込めないと課題がある。開発したモバイルアプリではカメラで撮影した画像と、事前に保存した画像の両方を入力することが可能だが、エンドユーザは主に端末のカメラから撮影した画像を使用すること

が想定される。そのため撮影環境を再現せずとも高精度な分類を行うことが可能な手法が今後の研究課題である。また、モバイルアプリでは、エンドユーザに対してアボカドの追熟の正しい知識を提供できるような機能の実装も行いたいと考えている。

参考文献

- [1] Katy A Cox, Tony K McGhie, Anne White, and Allan B Woolf. Skin colour and pigment changes during ripening of 'hass' avocado fruit. *Postharvest Biology and Technology*, 31(3):287–294, 2004.
- [2] Monzurul Islam, Khabou Wahid, and Anh Dinh. Assessment of ripening degree of avocado by electrical impedance spectroscopy and support vector machine. *Journal of Food Quality*, 2018:1–9, 11 2018.
- [3] Jhon Pinto, Hoover Rueda-Chacón, and Henry Arguello. Classification of hass avocado (persea americana mill) in terms of its ripening via hyperspectral images. *TecnoL*, 22:111–130, 08 2019.
- [4] Koyama K. Olivares Díaz E. et al. Cho, BH. Determination of "hass" avocado ripeness during storage based on smartphone image and machine learning model. *Food Bioprocess Technol*, 13:1579–1587, 7 2020.
- [5] M. Haggag, S. Abdelhay, A. Mecheter, S. Gowid, F. Musharavati, and S. Ghani. An intelligent hybrid experimental-based deep learning algorithm for tomato-sorting controllers. *IEEE Access*, 7:106890–106898, 2019.
- [6] Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Uproft, Tristan Perez, and Chris McCool. Deep-fruits: A fruit detection system using deep neural networks. *Sensors (Basel, Switzerland)*, 16(8), 2016.
- [7] Horea Mureşan and Mihai Oltean. Fruit recognition from images using deep learning. *arXiv e-prints*, page arXiv:1712.00580, Dec 2017.
- [8] José Naranjo-Torres, Marco Mora, Ruber Hernández-García, Ricardo J. Barrientos, Claudio Fredes, and Andres Valenzuela. A review of convolutional neural network applied to fruit image processing. *Applied Sciences*, 10(10), 2020.
- [9] Mahmut KAYA and Hasan Şakir BILGE. Deep metric learning: A survey. *Symmetry*, 11(9), 2019.
- [10] Pallabi Ghosh and Larry S. Davis. Understanding center loss based network for image retrieval with few training data. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, September 2018.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [12] Yandong Wen, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. A discriminative feature learning approach for deep face recognition. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 499–515, Cham, 2016. Springer International Publishing.