

# メロン画像特徴量の類似度を用いた等級判定技能習得支援システムの提案

小池 誠<sup>1</sup> 小川 晋<sup>2</sup> 峰野博史<sup>3</sup>

**概要**：温室メロンは、近年では年間約 1.8 万トン生産されており、国内市場では主に贈答に用いられるような高級果実として流通している。高級果実として生産されるメロンは、糖度などの味の品質だけでなく、形状や色合い、ネットの模様、傷の有無などの外観の品質も重要視されている。そのため、出荷時には生産者によって、果実の見栄えに関しても厳しいチェックが行われており、外観の品質を加味した等級の判定が行われている。この等級判定作業は、一般的には経験豊富な生産者による目視確認で行われている。しかし、経験に基づく作業であるため、等級判定のための目利き技能を習得するまでに多くの時間を要するという課題と、一度覚えた等級の判定基準であっても時間経過とともに自然とばらついてしまうという課題が存在する。そのため、現在は等級判定の基準を生産者間で統一するために、定期的な目揃い会を開催するなど、等級判定の一貫性担保と新人生産者の技能習得のために多くの時間と労力が割かれている。そこで本研究では、メロン表面画像を入力とした深層距離学習と説明可能な AI 技術による可視化を用いることにより、新規就農者向けの等級判定技能習得支援システムを提案する。そして、基礎的な実験を行い可視化の有用性を示した。

**キーワード**：農林水産，説明可能な AI，等級判定

## 1. はじめに

温室メロンは、近年では年間約 1.8 万トン生産されている。温室メロンは、国内市場では主に贈答用として用いられるなど、高級果実として広く知られており、温室メロン生産者も品質の高いメロンを出荷するために、多くの時間と労力を注いでいる。出荷時の等級判定作業もその一つで、温室メロンは果実の大きさや糖度などの味に関する品質に加え、果実の見え目といった外観の品位についても厳しくチェックされる。具体例を挙げると、果実の形状が整っていること、色合い、果実表面の網目の形成パターンの均一度、傷や汚れの有無などが評価され 4 つの等級に選別される。この等級判定作業は、一般的には果実一玉ごとに目視で行われるため、多くの労働時間が費やされている。そのため、等級判定作業の効率化や自動化が強く望まれている。また、目視で等級判定を行うため、生産者間で等級判定の基準がばらついてしまい、等級の一貫性が保たれないという課題も存在する。現状では、この課題に対処するため、生産者の間で等級の基準を照らし合わせる目揃い会を定期的に開催するなど、多くの時間と労力がかけられている。

本研究では、メロンの等級判定の効率化・自動化を目的としたメロン等級判定システム（以下、本システム）の提案を先行研究[1,2]で行ってきた。メロンの等級判定では、メロン表皮に現れる網目のパターンの識別が重要になるが、注目すべき網目パターンはメロン表皮に一様に分布しているわけではなく、メロン果実の成長過程の中で表皮の局所

的な部位に現れることが多い。また、メロン表皮に現れる網目パターンも、対象が自然物であるため多様なパターンを持っている。先行研究では、このような性質を持つメロン等級判定タスクに対応するため、等級間のメロン表皮画像における局所的な特徴量の類似度に着目し、深層距離学習を用いた等級判定モデルの構築と構築したモデルを組み込んだ等級判定システムのプロトタイプを開発した。そして、有用性に関する基礎実験を行った結果、約 84% の割合で熟練生産者が許容する等級判定を行うことができた。しかし、等級判定システムがメロンのどこを見て等級判定を行ったのかがユーザにとってブラックボックスであるため、システムの判定が熟練者の判定と異なっている場合、なぜシステムが判断を間違えたのか、又は、熟練者側で何か見落としている部分があるのではないかといった疑問が発生してしまい、システムが行う等級判定への信頼性が低下することが確認された。

そこで本論文では、説明可能な AI (XAI) 技術を用いて等級判定の根拠となる画像の特徴をユーザに分かりやすい形で提示することで、本システムが行う等級判定の信頼性向上を図る。また、等級判定の根拠となる等級ごとの特徴を視覚的に表示することによって、新規就農者などが各等級の特徴を覚えるための等級判定技能習得支援システムを提案する。

以降、2 章で関連研究について述べる。3 章で本システムを用いた等級判定の結果と判定根拠となった部位の提示方法について述べ、4 章では提示方法の有用性についての基礎実験を行った。5 章で結果の考察を行い、最後に 6 章で本論文をまとめる。

1 静岡大学大学院自然科学系教育部情報科学専攻  
Graduate School of Science and Technology, Shizuoka University

2 (株) 大和コンピューター  
DAIWA COMPUTER CO., LTD.

3 静岡大学大学院情報学領域 / グリーン科学技術研究所  
College of Informatics, Academic Institute, Shizuoka University /  
Research Institute of Green Science and Technology

## 2. 関連研究

### 2.1 説明可能な AI (XAI)

深層学習技術の急速な発展や深層学習に基づくシステムが様々な分野へ普及している一方で、深層学習に基づくモデルやシステムの透明性の欠如が課題となっており、深層学習を組み込んだシステムを導入する際の大きな障害となっていることが報告されている。説明可能な AI (XAI) [3, 4]は、深層学習の高い性能を維持したまま、より説明可能なモデルを生成する手法の構築を目的とした研究分野であり、モデル自体の透明性、ユーザ解釈のための可視化、説明の信頼性などに関する研究が盛んに行われている。

その中でも、ユーザ解釈のための可視化技術に注目する。ユーザ解釈のための可視化手法には、モデル特有の手法 (Model-specific) とモデルに依存しない手法 (Model-agnostic) が存在する。モデル特有の手法では、Zhou ら[5]による Class Activation Mapping (CAM)や、Selvaraju ら[6]による Grad-cam などがよく知られている。どちらも畳み込みニューラルネットワーク (CNN) を用いた分類タスクを対象とした分類根拠の可視化手法で、畳み込み層の特徴マップと出力層の重み、又は、出力層の出力に対する逆伝搬により求められる特徴マップの勾配を用いて、分類結果に大きく寄与した画像をヒートマップとして表示。また、Chen ら[7]は Grad-cam を拡張し、トリプレットなどに代表される深層距離学習においても Grad-cam を適用する Adapting Grad-CAM を提案すると共に、テスト時に特徴ベクトルの類似度に基づく最近傍法を適用することで、ヒートマップ画像生成の高速化法について報告している。一方、モデルに依存しない手法では、Ribeiro ら[8]の LIME に代表されるような、説明対象のモデルをブラックボックスとして扱い、モデルの入力に摂動を加えた場合に出力がどのような影響を受けるのかを分析することで、対象モデルの説明を試みる手法が提案されている。LIME は、説明対象モデルの入力と出力の関係を、任意のデータに摂動を与えることで、そのデータの周辺における対象モデルの入力と出力の関係を、別の線形回帰モデルのような単純で説明可能な代替モデルで表現することによって、対象モデルの局所的な説明を試みる。

本システムでも、Zhou らの CAM で提案された Activation Mapping の考え方を基にした手法を用い、メロンの等級判定に大きく寄与した部位の可視化を行う。そして、可視化部位を等級判定結果と同時に表示することで、新人に対する等級判定技能習得のサポートを試みる。

### 2.2 深層距離学習

深層距離学習[9]は、入力データ間の類似性、又は、非類似性を確立するために深層学習のアーキテクチャを用いる手法である。入力データを特徴空間上のベクトルに埋め込

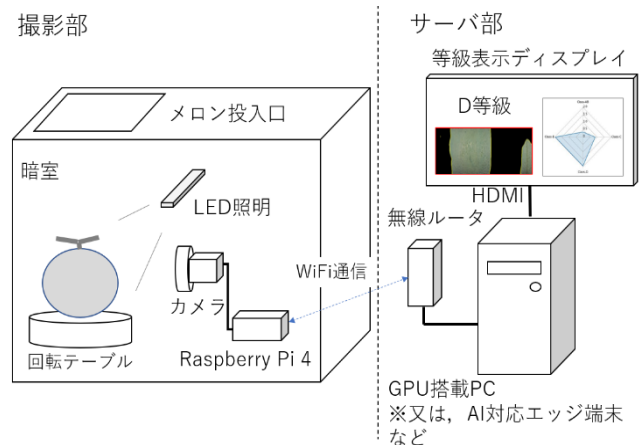


図1：等級判定技能習得支援システム

む際、同じラベルのデータは特徴空間上で近い距離になるように、逆に、異なるラベルのデータは遠い距離になるように学習が進む。主に、画像を用いた顔認識・再認識タスクや異常検知タスクなどで用いられており、一般的な他クラス分類モデルを用いる場合と比較し、未知のデータに対する分離性や特徴ベクトルの類似性をスコアとして、別タスクで活用できるといったメリットがある。

本システムでは、メロンの網目の特徴を表すために、距離学習手法の1つである ArcFace[10]を用いる。深層距離学習を用い、等級ごとのメロン網目の画像を画像特徴の類似度による特徴ベクトルに変換する。そして、特徴ベクトルのコサイン類似度を用いて等級を分類する。また、判定対象となったメロンに対し、各等級にどの程度似ているかを提示することで等級判定技能習得のサポートを試みる。

## 3. メロン等級判定技能習得支援システム

図1に、本システムの構成を示す。本システムは、撮影部とサーバ部で構成されており、撮影部ではカメラを用いてメロンの360°全周画像を取得し、サーバ部では撮影部で取得した全周映像の前処理、等級判定モデルを用いた等級判定、等級判定の結果と判定根拠の提示を行う。なお、本システムの詳細については、先行研究[2]で説明しているため、本論文では主に先行研究から拡張した技能支援システムに関する部分について述べる。

### 3.1 メロンの等級について

本システムで対象となるメロンは、静岡県内で温室栽培されたアールスメロン系品種で、2月から11月にかけて収穫されたものである。等級判定は、熟練生産者によって行われており、品質の高い順に、A, B, C, Dの4等級に選別される。各等級の判定基準について、生産者から聞き取りを行った結果を示す。

- ・ A 等級：高品質  
網目の密度が均一で果実の形状に歪みが少ない
- ・ B 等級：中品質  
網目の乱れが目立つものの果実の形状に歪みがない
- ・ C 等級：低品質  
果実形状に扁平，面長，凸部など歪みがある
- ・ D 等級：低品質  
果実に傷，汚れ，黒カビ，著しい網目の乱れがある

等級付けの優先順位は，D>C>B>A 等級の順となり，品質の低い等級の特徴が優先される．例えば，C と D の特徴を持つメロンは，D 等級と判定される．なお，輪郭の特徴をもって判断される C 等級と，汚れや傷などの特徴をもって判断される D 等級は比較的判断が容易であり，網目の特徴をもって判断される A, B 等級の判断が難しいとされている．網目の特徴によって判断される A, B, D 等級のサンプルを図 2 に示す．図 2 を見ると，メロン個体ごとに網目のパターンは多様であり，網目の大きさやばらつき具合などの定量化が難しく，熟練者であっても等級判定に重要な網目の特徴を言語化し新人へ伝えることが難しいことが分かる．

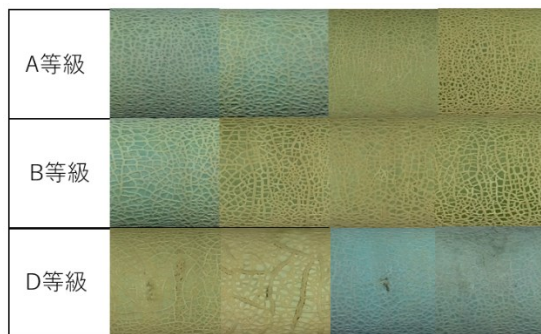


図 2：等級別の網目パターンの特徴

### 3.2 撮影とデータの前処理

本システムでは，撮影部でメロンの 360° 全周映像データを取得しサーバ部へ転送する．サーバ部では，データの前処理として，映像データから画像処理を用いて，輪郭画像と網目画像を生成する．前処理後の網目画像の例を図 3 に示す．1 玉のメロンからパノラマ画像を生成する要領で，1 枚の網目画像を生成する．

### 3.3 ArcFace を用いた等級判定モデル

先行研究[2]で述べた等級判定モデルの学習アーキテクチャの概要を図 4 に示す．エンコーダには ResNet50[11]を使用し，MC-Loss[12]を用いて網目画像から A, B, D の 3 等級を表す網目特徴ベクトルへの変換を学習する．輪郭画像は，画像処理で算出した統計量による輪郭特徴ベクトルへ埋め込みを行い，網目特徴ベクトルと輪郭特徴ベクトル

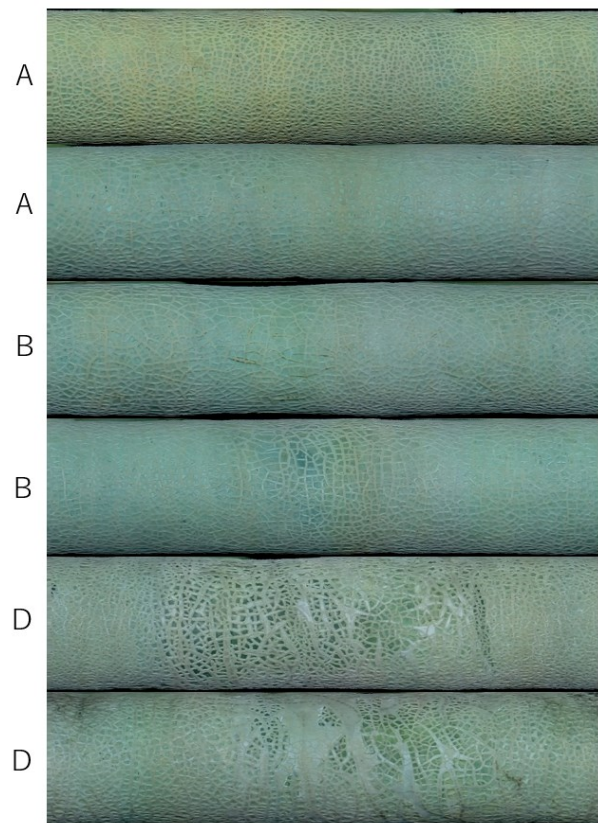


図 3：等級別の網目画像の例

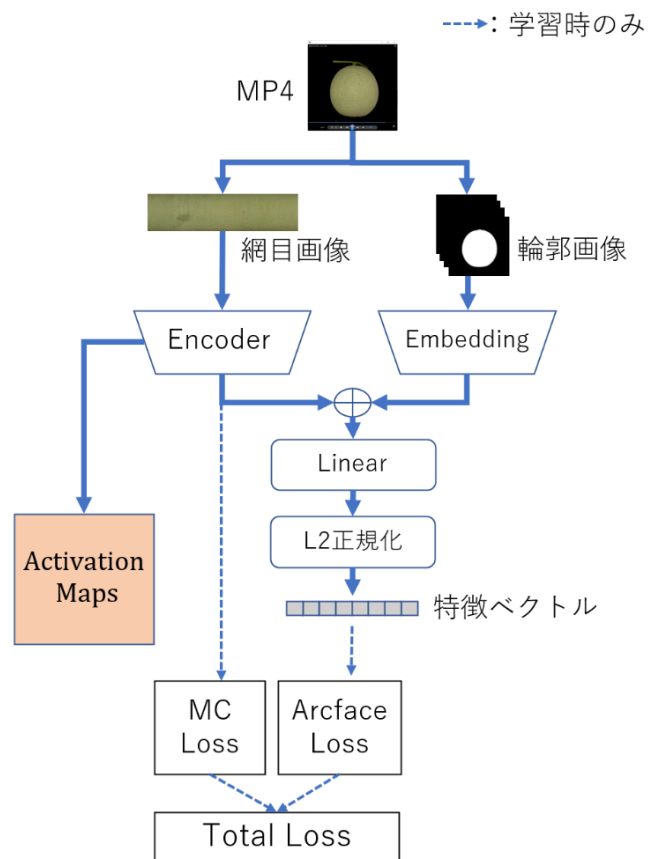


図 4：等級判定モデルのアーキテクチャ概要

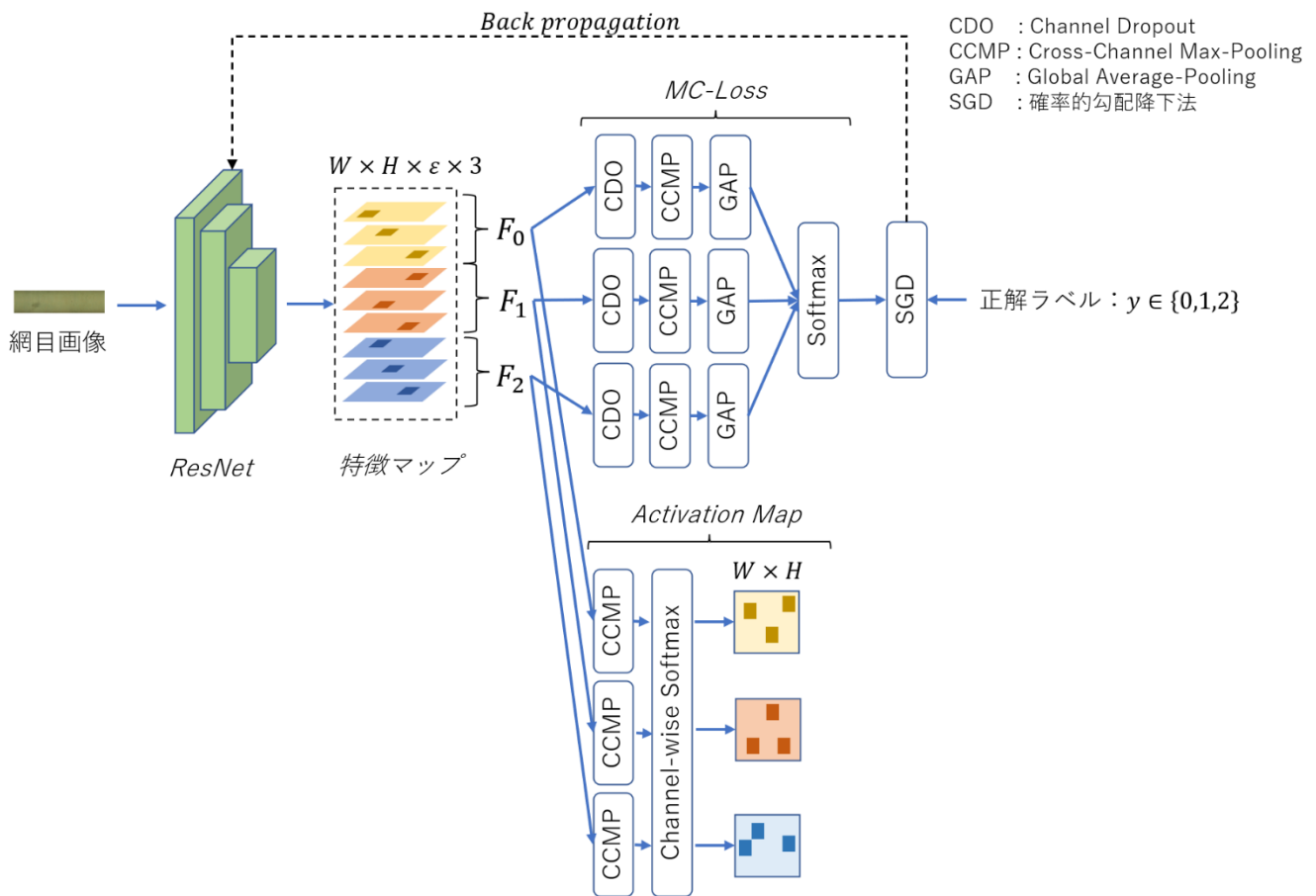


図 5 : Activation Map の生成手順

を連結し、ArcFace を用いて 4 等級の分類を学習する。

### 3.4 Activation Map を用いた特徴部位の可視化

等級判定に寄与した部位を表す Activation Map の生成方法について述べる。図 5 に Activation Map 生成の過程を示す。Activation Map は、エンコーダの最終畳み込み層が出力する特徴マップを使って生成する。この時、特徴マップは MC-Loss を用いることで、各等級の判定に有効な特徴が特徴マップの等級別チャンネル  $\epsilon$  上に現れるよう学習される。よって、特徴マップの高さを  $H$ 、幅を  $W$ 、分類する等級数を 3 とすると、等級ごとの特徴マップは、 $F_i \in \mathbb{R}^{\epsilon \times W \times H}, i = 0, 1, 2$  と表すことができる。各特徴マップは MC-Loss により、ランダムに選択したチャンネルをドロップアウトする Channel Dropout (CDO)、チャンネル方向にマックスプーリングを取る Cross-Channel Max-Pooling (CCMP)、チャンネルごとの平均値を算出することで、特徴マップの平坦化を行う Global Average-Pooling (GAP) が適用された後、一般的な多クラス分類と同様に Softmax により各クラスの確率分布に変換され、確率的勾配降下法を用いて ResNet の学習が行われる。これら一連の処理により、ResNet は入力される網目画像のラベルを表す画像特徴の総量が大きくなるように特徴抽出を行うエンコーダとして学習が進む。Activation

Map は、この等級ごとの特徴マップ  $F_i$  で CCMP を適用し、チャンネル方向の Softmax を取ることで生成する、言い換えると、Activation Map は、画像空間  $W \times H$  上でピクセルごとにどの等級であるかを確率で表すマップであると言える。

本システムの Activation Map は、Zhou ら[5]の Class Activation Mapping (CAM) と基本的なアプローチは同等である。しかし、CAM では畳み込み層の特徴マップ出力の後段に、GAP と全結合層によるクラス分類器が配置されるため、全結合層の重みによって特徴マップの値がリスケールされる。よって、CAM では Activation Map の生成に全結合層の重みを考慮する必要があった。それに対し、本システムは MC-Loss を用いることで特徴マップの出力の後段にリスケールを行う処理が不要となる。よって、特徴マップから直接 Activation Map を生成できるため、計算コスト削減というメリットがある。

### 3.5 画像特徴の類似度の可視化

本システムは、等級判定モデルに ArcFace を用いることで、メロン画像の特徴を表す特徴ベクトルを角度距離空間へ埋め込んでいる。ArcFace では、同じ等級の特徴ベクトルが角度距離空間上で近い位置に、異なる等級の特徴ベクトルは遠い位置へ配置されるように学習が進む。よって、学

学習済みモデルが出力する特徴ベクトル間の類似度を、コサイン類似度により表すことができる。

本システムの等級判定では、特徴ベクトルの類似度を用いて等級判定を行う。まず、学習データを用いて各等級の平均特徴ベクトル  $REF_i, i = 0, 1, 2, 3$  を求める。運用時は、事前に求めておいた各等級の平均特徴ベクトルと判定対象となるメロンの特徴ベクトル  $F$  との類似度  $S_i, i = 0, 1, 2, 3$  を求め、最も類似度が高い等級をそのメロンの等級  $y'$  と推定する。

$$S_i = \frac{F \cdot REF_i}{\|F\| \|REF_i\|} \quad (1)$$

$$y' = \underset{i \in \{0, 1, 2, 3\}}{\operatorname{argmax}} S_i \quad (2)$$

等級判定の結果は、ディスプレイに表示する。その際、推定された等級  $y'$  のみを表示するのではなく、各等級との類似度  $S_i$  も併せて表示する。本システムの等級判定のプロセスをユーザに開示することにより、本システムが行った判定の透明性向上を図る。具体的には、本システムがあるメロンの等級判定を行った際、提示した等級が他の等級と比べ極めて類似度が高かったのか、又は、他の等級と僅差であったのかをユーザが認識できるようにする。このような表示は、距離学習を用いない一般的な多クラス分類モデルにおいても Softmax の結果を指標することで実現可能であるが、Softmax の結果は合計すると 1.0 になる値に正規化されてしまうため、どの等級にも似ていないという場合の提示ができない。それに対し、本システムのように類似度を指標とした場合は、Softmax のような制約はないため、どの等級とも似ていないという表現が可能となり、学習用データには含まれなかった特徴のメロンも扱うことができる。

## 4. 基礎実験

### 4.1 データセット

基礎実験で使用したデータセットは、各等級 122 玉のメロンの全周映像データである。データの取得は、2021 年 2 月から 2021 年 12 月にわたり 18 回実施され、等級のアノテーションは 1 名の熟練生産者によって行われている。

### 4.2 実験条件

実験で使用する等級判定モデルは、等級ごとにランダムに選択した 92 玉を学習用データ、残り 30 玉をテストデータとし、エンコーダには imagenet で事前学習された ResNet50 を使い、モーメントム=0.9 に設定した SDG、バッチサイズ=32、学習率=0.001、エポック数=250 として学習を行った。学習の結果、3 回の交差検証で正答率 82.1% となる等級判定モデルを構築することができた。これを可視化実験に用いる学習済みモデルとした。

表 1: 可視化手法別の mIoU と処理時間

	LIME	Grad-CAM	提案手法
mIoU	0.30	0.36	<b>0.62</b>
処理時間 (ms)	6329	36	<b>35</b>

実験に使用した PC は、CPU: Intel Core i7-6700K@4.0GHz、主記憶メモリ: 64GB、GPU: NVIDIA GeForce GTX1080 を搭載したデスクトップ PC である。

### 4.3 Activation Map の有用性の検証

等級判定モデルの判定に大きく寄与した網目画像の部位の可視化実験を行った。学習済みモデルを用いてテストデータに対する Activation Map の生成を行った。比較対象として、Adapting Grad-CAM と LIME による可視化画像の生成も行った。Adapting Grad-CAM では、ResNet が出力する特徴マップに GAP を適用して特徴ベクトルとし、テスト時の最近傍法ではこの特徴ベクトルを用いて重みの探索を行った。LIME では、スーパーピクセルを求める手法に QuickShift 法[13]を用いた。周辺データのサンプリング数を 1000 としてマップの生成を行った。

等級を表す画像特徴部位を可視化できているかを評価する指標は mIoU[6]を用いた。まず、熟練生産者に等級を表している画像部位をアノテーションしたマスク画像を作成してもらい、これを正解マップ (図 6 の Ground truth) とした。次に、A, B, D 等級の 90 玉のテストデータを用いて、各手法で生成した Activation Map をしきい値 0.5 で 2 値化したマスク画像を生成し、正解マップとの mIoU を求めた。

実験結果として、テストデータに対する mIoU と可視化画像 1 枚を生成するために費やした処理時間を表 1 に示す。比較対象とした 3 つ手法の中で、本システムで用いた手法の mIoU が 0.62 と最も高い値であった。また、可視化画像生成にかかる時間も 35ms と最も短い時間で生成できることが確認できた。また、実際に生成された可視化画像を図 6 に示す。提案手法の可視化結果が、他の手法と比較し、正解マップの白くアノテーションされた部位と一致していることが確認できる。

### 4.4 類似度の提示による有用性の確認

3.5 節で示した等級ごとの平均特徴ベクトルと等級判定対象のメロンの特徴ベクトルとの類似度を等級判定の結果と同時にユーザへ提示することが、メロンの特徴学習に有用かを確認した。まず、テストデータより各等級 20 玉を選択し、等級判定とコサイン類似度の算出を行った。その中で、コサイン類似度が 1.0 以上となる等級が複数あるメロンを選び個数を計測した。複数の等級に同時に類似しているメロンは、新人生産者にとって判断が難しいメロンと予想される。次に、選択したメロンを使い、本システムによ

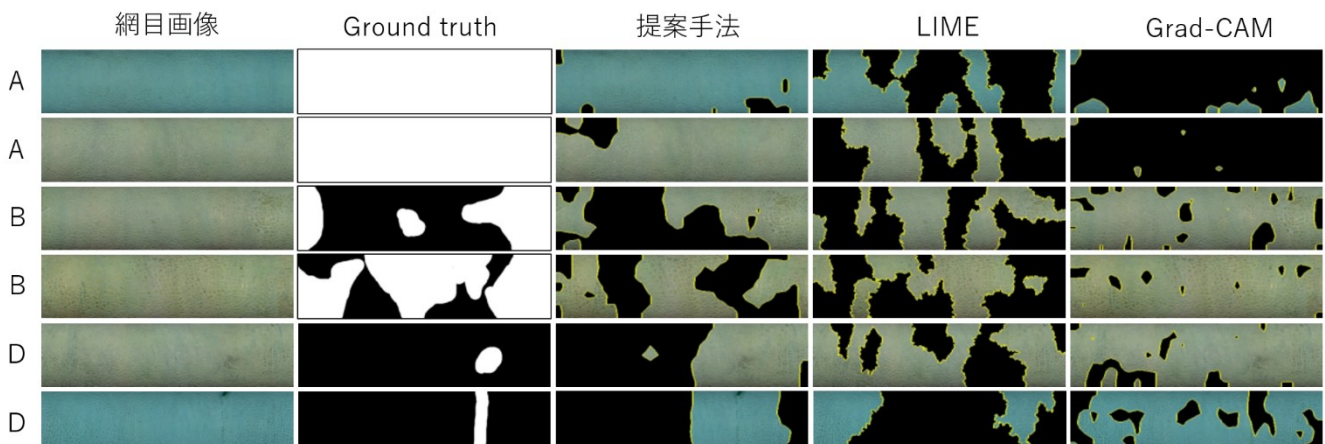


図6：等級別の各手法による可視化画像例

る等級判定結果と類似度を図7に示すレーダーチャート形式で提示し、本システムの判定結果について熟練生産者2名に次の3項目で回答してもらった。

1. 判定が一致している
2. 一致していないがレーダーチャートは妥当である
3. 明らかに間違っている（見間違いである）

結果を表2に示す。複数の特徴を持つメロンの数は、85玉であり全体の53.1%であった。中でもC等級は、全て複数の特徴を持つという結果であった。これらのメロンに対する熟練者の回答を集計した結果を、同じく表2に示す。この内、熟練者の回答が1、又は、2であったものは、表示したレーダーチャートが目視で確認したメロンの特徴を正しく反映したものであったと言える。そこで、回答が1、又は、2であった数から妥当なチャートの割合を算出した結果、全体の97.6%であった。

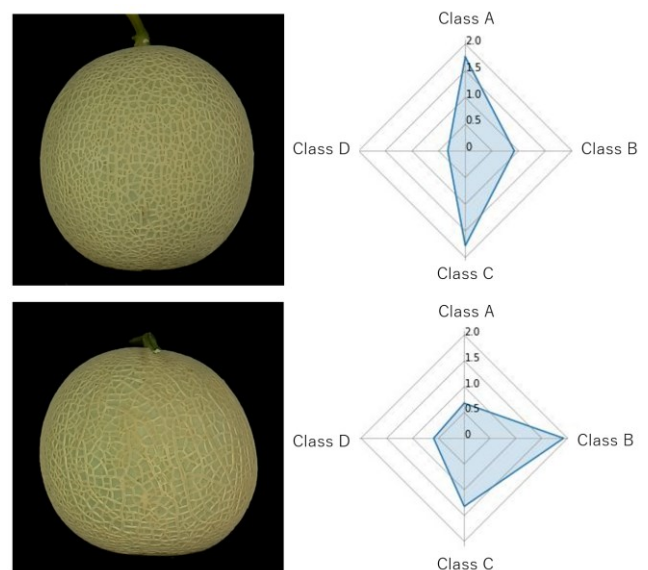


図7：類似度のレーダーチャート表示例

## 5. 考察

表1と図6より、MC-Lossを用いたActivation Mapが、比較対象よりも、網目の特徴部位を正しく表すことが確認できた。提案手法が比較手法よりも良い結果となった要因の1つに、学習データ数が少ないことが考えられる。各等級92玉程度のデータ数であるため、一般的な転移学習を用いたResNetのファインチューニングでは過学習となり、等級に有効な特徴抽出や全結合層の重みを学習できなかった可能性がある。それに対し、MC-Lossを用いる場合は、等級判定に有効な局所的な特徴のみを抽出するよう学習が進む。通常のResNetのような、特徴抽出で様々な特徴を多く（特徴マップのサイズ=7×7×256）抽出し後段の全結合層で分類に有効な重み付けを行うアプローチよりも、MC-Lossのような、特徴抽出の段階で等級判定に有効な特徴のみ（特徴マップのサイズ=7×7×15）抽出するアプローチがActivation Mapを生成する上でも有効だと考えられる。な

表2：特徴を複数持つメロンの個数と熟練者の回答の集計結果

	複数特徴有		熟練者の回答			妥当なチャート割合(%)
	数	割合(%)	1	2	3	
A	17	42.5	10	7	0	100.0
B	5	12.5	3	2	0	100
C	40	100.0	29	8	2	94.9
D	23	57.5	22	1	0	100
Total	85	53.1	64	18	2	97.6

お、MC-Lossを用いた網目画像による3等級判定モデルを学習して汎化性能を評価したところ、正答率は75.8%であった。それに対し、通常のResNet50で学習を行ったモデルでは正答率67.9%であり、MC-Lossを用いたモデルの方が、等級判定に有効な特徴を抽出できていると考えられる。

レーダーチャート形式で類似度をユーザに提示する実験では、C 等級、D 等級のメロンは複数の等級との類似度が高い傾向にあることが分かった。C 等級のメロンは、唯一輪郭の歪みの特徴を持っており、加えて網目の特徴も併せ持つため、複数の類似度が高くなるのは自然であるといえる。また、D 等級のメロンは、傷や汚れなど表皮の極一部にある特徴と、それ以外の大部分を占める特徴とが複数検出されていると考えられる。また、熟練者の回答を集計した結果より、目視で確認した等級の特徴に対し、レーダーチャートの内容が 97.6%の割合で妥当であるという結果を得ることができた。これは単純に等級判定結果のみを表示する場合と比較し、よりユーザに対象メロンの詳細な特徴の情報を提示できる可能性を示すことができたと考えられる。これにより、新人が類似度情報を参考に、メロンの特徴を学習するような支援システムが考えられる。

## 6. まとめ

本研究では、Activation Map による等級判定に寄与した部位の可視化と距離学習を用いた類似度指標の提示を行うことで、システムの等級判定を補足する情報を付与し、新人の等級判定技能習得を支援する機能拡張について検討した。MC-Loss を用いた Activation Map により、学習データが少ない場合であっても、有用な特徴部位の可視化を行えることを確認できた。また、メロンの等級別類似度をレーダーチャート形式でユーザに提示することによって、メロンにどの等級の特徴が含まれているかを効果的に伝達できる可能性を示すことができた。これにより、システムが判定した等級のみを提示する場合と比較し、新人がメロンのどの部位に注目すべきかをサポートするような等級判定技能習得支援システムの実現が期待できる。

今後の研究では、実際に新人生産者を対象に、支援情報を提示することにより等級判定の技能習得がどの程度容易になるのか、本システムの支援効果の評価を進めていく。

**謝辞** 本研究の一部は、JST 創発的研究支援事業 (JPMJFR201B) ならびに東北大学電気通信研究所における共同プロジェクト研究の支援を受けたものである。また、データセットならびに有意義なディスカッションの機会をいただいた (株) 大和コンピューターの皆様に深い感謝の意を表す。

## 参考文献

- [1] 小池誠, 平原健太郎, 渡辺隆一, et al.: 半教師あり学習を用いた果実画像からのメロン等級判定システムの検討, 研究報告モバイルコンピューティングと新社会システム (MBL), Vol.13, pp.1-8, 2021.
- [2] 小池誠, 平原健太郎, 渡辺隆一, et al.: 画像局所特徴の類似度を用いたメロン等級判定システムの開発, CDS トランザ

- クシオン, 2022 (投稿中) .
- [3] Adai,A., Berrada,M.: Peeking inside the black-box: a survey on explainable artificial intelligence(XAI), IEEE access, Vol.6, pp.52138-52160, 2018.
  - [4] 原聡: 私のブックマーク「説明可能 A I」, 人工知能, Vol.34, No.4, pp.577-582, 2019.
  - [5] Zhou,B., Khosla,A., Lapedriza,A., et al.: Learning deep features for discriminative localization, In Proc. Of the IEEE conference on computer vision and pattern recognition, pp.2921-2929, 2016.
  - [6] Selvaraju,R.R., Cogswell,M., Das,A., et al.: Grad-cam: Visual explanations from deep networks via gradient-based localization, In Proc. of the IEEE international conference on computer vision, pp.618-626, 2017.
  - [7] Chen,L., Chen,J., Hajimirasadeghi,H., Mori,G.: Adapting grad-cam for embedding networks, In Proc. of the IEEE/CVF Winter Conference of Applications of Computer Vision, pp.2794-2803, 2020.
  - [8] Ribeiro,M.T., Singh,S., Guestrin,C.: Why should I trust you? Explaining the predictions of any classifier, In Proc. of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp.1135-1144, 2016.
  - [9] Kaya,M., Bilge,H.S.: Deep metric learning: A survey, Symmetry, Vo.11, No.9, pp.1066, 2019.
  - [10] Deng,J., Guo,J., Xue,N., Zafeiriou,S.: ArcFace: Additive Angular Margin Loss for Deep Face Recognition, In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.4690-4699, 2019.
  - [11] He,K., Zhang,X., Ren,S., Sun,J.: Deep residual learning for image recognition, In Proc. of the IEEE conference on computer vision and pattern recognition, pp.770-778,2016.
  - [12] Chang,D., Ding,Y., Xie,J., et al.: The Devil is in the Channels: Mutual-Channel Loss for Fine-Grained Image Classification, IEEE Transactions on Image Processing, Vol.29, pp.4683-4695, 2020.
  - [13] Vedaldi,A. Soatto,S.: Quick shift and kernel methods for mode seeking, In European conference on computer vision, Springer, Berlin, Heidelberg, pp.705-718, 2008.