

# 修士論文

## 農作物の自動選別に向けた多視点 データセットと品質分類モデルの構築

令和7年1月23日提出

指導教員 山崎 俊彦 教授

東京大学大学院情報理工学系研究科電子情報学専攻

48-236442

水谷航悠



# 概要

日本の農業現場では、高齢化と若年層人口の減少に伴う労働力不足が深刻化しており、特に作物の選別工程が生産コスト全体の約 1 割を占める大きなボトルネックとなっている。本研究では、この選別作業を自動化することを目的として、以下の 2 つのアプローチから作物選別精度の向上と研究分野への貢献を図った。

まず、二次元画像と三次元点群を組み合わせるマルチモーダルモデルの有効性を探るため、みかんを対象に実験を行った。200 個のみかんを撮影・3D スキャンし、データセットを作成した。ResNet を用いた画像モデルと PointNet を用いた点群モデルを学習させたところ、画像単体では 87.5% の精度を示す一方、点群単体は 65.0% に留まった。しかし、両者を同時に学習するマルチモーダルモデルでは 95.0% の高精度を記録し、二次元と三次元の情報を融合することが外観品質選別の大幅な精度向上につながることを示した。一方で、三次元データの取得に際して高精度な 3D スキャナーが必要となるなど、設備投資や運用面でのコストが大きい点も課題として浮上した。

次に、多視点から作物を撮影した大規模な公開データセットが存在しないことから、複数方向から農作物を高速に撮影できる専用機械を設計・製作した。撮影対象はレモン・みかん・じゃがいもの 3 種類であり、4〜5 方向から得られた数千枚規模の画像を収集・整備した。サイズや等級のラベルを自動的に付与し、単一視点では把握しづらい形状の特徴や微妙な傷・色むら、またサイズを総合的に学習可能なデータセットを形成した。さらに、こうして得られたデータセットを用いて深層分類モデル (ResNet-18) を学習させ、3 等級に対する分類性能を検証した結果、レモンで 85.3%、みかんで 89.7%、じゃがいもで 84.3% の精度をそれぞれ達成した。

インタビュー調査からは、農家が既存の AI 選果機の導入を躊躇する理由として、「設備投資の困難さ」「多様な選別基準への対応」「多品目同時処理の必要性」が挙げられた。すなわち、高額な AI 選別機は小規模農家にとって導入が難しく、さらに地域や品種によって異なる基準を柔軟に設定できる仕組みが欠かせない。本研究で構築した大規模・多視点データセットや、マルチモーダルモデルによる高精度選別手法は、こうした現場ニーズに応える基盤技術となり得るが、システム全体としては低コスト化や UI 設計、転移学習を用いた多品目対応など、社会実装に向けた総合的な工夫が必要である。



# 目次

第 1 章	序論	5
1.1	背景と問題設定 . . . . .	5
1.2	本論文の構成 . . . . .	9
第 2 章	関連研究	11
第 3 章	3 次元点群情報を考慮した農作物の外観品質選別	14
3.1	本研究の背景 . . . . .	14
3.2	関連研究 . . . . .	16
3.3	提案手法 . . . . .	20
3.4	実験 . . . . .	21
3.5	結果と考察 . . . . .	23
3.6	未来の研究方向性 . . . . .	25
3.7	結論 . . . . .	26
第 4 章	農作物選別作業の自動化に向けたデータセット構築と品質分類モデルの開発	28
4.1	本研究の背景 . . . . .	28
4.2	関連研究 . . . . .	29
4.3	データセット作成 . . . . .	29
4.3.1	撮影機械の設計と開発 . . . . .	29
4.3.2	データ収集の協力体制 . . . . .	31
4.3.3	データセットの構成 . . . . .	33
4.3.4	データセットの価値と意義 . . . . .	34
4.4	実験 . . . . .	36
4.5	実験結果と考察 . . . . .	37
4.6	未来の研究方向性 . . . . .	38
4.7	結論 . . . . .	39
第 5 章	自動 AI 選果システム実現に向けた現在の課題と未来の研究方向性	41
5.1	インタビューを通じて判明した課題 . . . . .	41



5.2	実地導入のために追加で必要な要件 . . . . .	42
5.3	今後の展望と研究方向性 . . . . .	42
第 6 章	結論	44
参考文献		46
発表文献		48



# 第 1 章

## 序論

### 1.1 背景と問題設定

現在、日本国内の地方においては若年層の人口が減少しており、農業分野における労働力不足が顕著な問題となっている。その影響で、基幹的農業就業人口は 2010 年の 205 万人から 2015 年には 175 万人、2020 年には 136 万人と減少を続けている (図 1.1) [1]。一方で、農家一戸あたりの経営耕作面積は平均で約 1.5 倍に拡大している [1]。また、農業従事者の平均年齢が 68 歳を超えるなど高齢化が進んでおり [2]、生産規模の維持や拡大のためには、農業プロセスの省人化や機械化が急務である。

農業プロセスは、定植→施水施肥→収穫→選別→出荷に大きく分割できる。十数回にわたる農業施設へのインタビュー及び訪問を通じて、国内の農業サイクルの中で、作物の選別工程が省人化の一つのボトルネックであることが明らかになった。選別された作物は、未選別の作物に比べて一般に高単価で取引されるため、選別工程は、農業経営体として必須のプロセスである。青果物の出荷規格は、取引を円滑に進め、有利な販売を実現するために産地ごとに設定されている。しかし、産地間の競争やブランド化により規格の過度な細分化が進んでおり、これが選別の複雑化を招いている。さらに、海外と比べて日本国内の選別基準が厳しく、選別される品種も多様であるため、農協や農業法人はアルバイトの確保や施設運営費に苦慮している状況が見受けられる。

日本の農作物選別には、「個選」と「共選」の二つの主要な方法が存在する。個選 (個別選果) は、農家自身が収穫した作物を自ら選別し、出荷する方法である。個選では、自家用の簡易な選別機械や手作業によって品質やサイズが選別されることが多い。個選農家は作物を直接市場や消費者に届けるため、農家にとって品質管理の自由度が高い反面、選別作業の負担が大きい。共選 (共同選果) は、複数の農家が共同で収穫物を集め、農協 (JA) などの施設で選別を行う方法である。共選は効率的に選別を進めるための仕組みであるが、農協に出荷する前段階で一定の品質基準を満たすよう、農家が事前に選別を行う必要がある場合が多い。

これらの選別方法のいずれにおいても、選別工程を完全に他者に委託することは難しく、農家自身が一定の労力を割くことになる。特に選別のピーク時期には、個選・共選を問わず、臨時のアルバイトを雇用して対応する農家が多いのが現状である。人手不足や高齢化の進む地方では労働



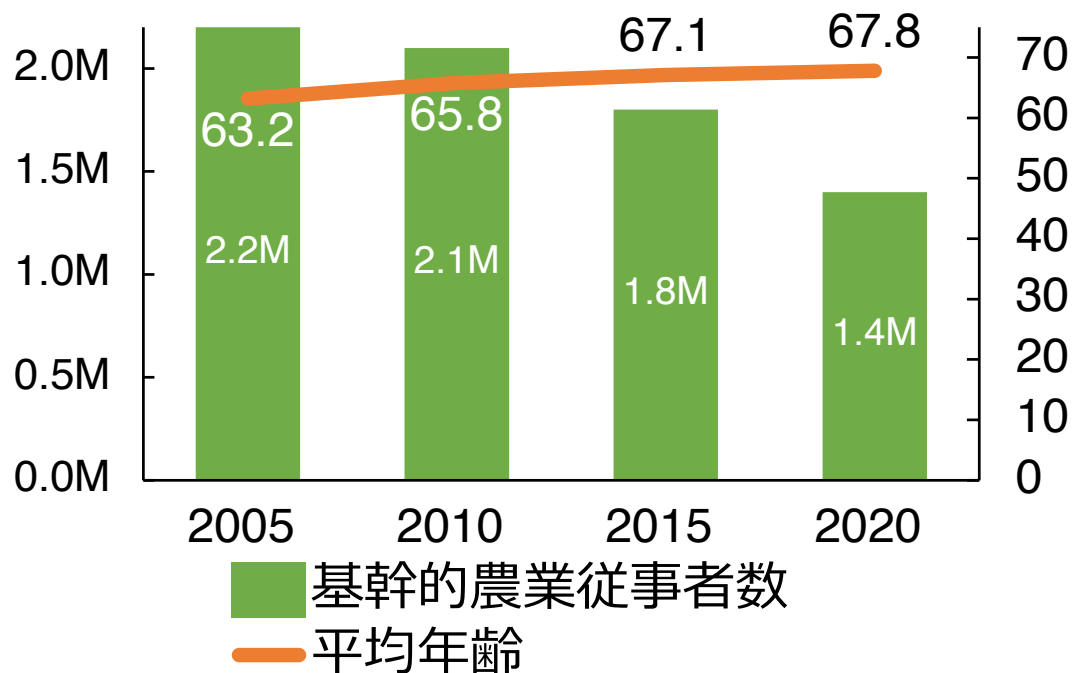


図 1.1: 日本の基幹的農業従事者の数と平均年齢.

力の確保も難しくなっており、これが農業経営における大きな負担となっている。実際、国内の作物選別工程には年間約 3300 億円が投じられており、これは日本の農業経営費の約 1 割に相当する [3].

農作物の選別工程は主にサイズ・重量選別と、等級選別の 2 段階に分けられる (図 1.2). 多くの場合、サイズ・重量選別は機械化されているが、等級選別は手作業で行われている。

**サイズ・重量選別 (規格選別) :** 第一段階では、サイズや重量に基づいて作物を選別する。このプロセスは物理的な指標に基づくため、比較的機械化が進んでおり、多くの農業施設でサイズ選果機が導入されている。図 1.3 に実際のサイズ選果機の写真を示す。サイズ選果機は、回転するローラーや振動式のトレイを用いて、作物を異なるサイズのトレイやコンテナに仕分ける。重量センサーを搭載した装置もあり、重量に基づく分類が可能である。このような機械化により、第一段階の選別作業は高い精度で効率化されている。

**等級選別 (品質選別) :** 第二段階の等級選別は、作物の形状、色、損傷、変形、緑化など、多岐にわたる基準に基づいて行われる。このプロセスは複雑であり、単純な物理的指標では判別が困難なため、手作業に依存しているのが現状である。図 1.4 には、等級選別のために農業施設で見本として置かれているレモンの写真を示す。選果者はこの見本と、次のような選果基準を参考に、色や形、傷を見ながら選果を行う (図 1.5).





図 1.2: 農業サイクルにおける選別プロセスの流れ.

- 1 級品の下限
  - － 傷の程度は人差し指 2 本で隠れる大きさまで
  - － 果実の突起は滑らかな程度で 2 カ所まで
  - － 黒点病は目立たない程度
  - － 縦径が横径よりも長い
- 2 級品の下限
  - － 人差し指 2 本で隠れる大きさの傷が 2 カ所まで
  - － 果実の突起は滑らかな程度で 5 カ所まで
  - － 黒点病はまばら、表面積の 3 分の 1 まで
  - － 丸形果は酔いが扁平果・奇形果は NG
- 規格外となるもの
  - － 各等級の規格に当てはまらないもの
  - － 病気、生傷果、遅れ花の果実

また、当然等級選別の基準は作物ごとに異なる。例えば、キャベツでは汚れや虫食いの有無、きゅうりでは曲がり具合、みかんではへた部分の張り、イチゴや桃では傷の有無など、選別にはそれぞれ特有の判断基準が求められる。そのため、選別者は作物ごとに特化したスキルを習得する必要があり、この複雑性が等級選別の自動化を困難にしている。結果として、等級選別の多くは依然として手作業に依存している。品質選別業務では、作物 1 つにつき 2～3 秒の業務スピードを保ちながら、1 日に数十トンの作物を選別する。十分なスキルを獲得するには一定の訓練と高い集中力が必要であり、新人やアルバイトがラインに入ると選別先がわからず、頻繁にラインを止めて熟練者に尋ねる様子も見られる。また、選別基準は指導員によって多少異なるため、選別者間での基準の共有や施設間の人材移動も困難である。各選別施設では過疎化や高齢化に伴い業務効率が低





図 1.3: サイズ選果機 (ヤマシシ果樹園にて撮影).

下し、ベテラン選別者の引退や慢性的な人手不足が問題となっており、今後、選別人員の確保はさらに難しくなると予想される。

選別作業の課題を克服し、自動化を実現するためには、コンピュータビジョン技術を活用した高精度な分類モデルの開発によって、人が目で見て行っていた作業を代替することが不可欠である。コンピュータビジョン技術を活用した高精度な分類モデルを開発するうえで、まず重要となるのは、実際の生産現場の多様な選別基準や複雑な外観特徴を網羅した大規模データセットの整備である。そこで本研究では、レモン・みかん・じゃがいもを対象に複数方向から撮影できる撮影機械を独自に開発し、多視点画像データを大規模に収集する。また、作物を 3D スキャンして画像データだけでなく 3D 点群データを収集することで、作物の立体的な形状も含めた等級選別が可能となる。これらのデータを用いて、二次元画像のみを入力とする従来型モデルと、画像＋点群を組み合わせたマルチモーダルモデルの両方を構築し、精度比較を行う。加えて、構築した多視点の大規模データセットでの精度検証も行う。さらに、農協や農家へのインタビュー調査を行い、AI 選別導入におけるコストや多品目対応といった現場ニーズを分析することで、研究の成果を具体的な運用へとつなぐシステム要件を整理する。





図 1.4: 等級選別の見本として置かれたレモン (JAひろしま くれ選果場にて撮影).

本研究は最終的に、大規模・多視点データセットや3D点群マルチモーダルモデルを通して、農作物等級選別の自動化を一步進めることを目指す。その結果として、労働力不足の深刻化に直面する国内農業において、省力化と効率化に寄与するAI選別システムの基盤技術を提供し、作物の生産性向上と持続可能な農業経営に貢献することが期待される。

## 1.2 本論文の構成

本論文は以下の構成で進める。

- 第1章：研究背景と問題設定を示し、国内農業における選別作業の重要性と課題を概観する。
- 第2章：農作物選別に関わる関連研究を幅広く整理し、AI技術を用いた既存の取り組みを概観する。
- 第3章：3次元点群情報を活用した外観品質選別手法として、PointNetと画像分類モデルを組み合わせた実験を行い、その有効性と課題を論じる。
- 第4章：多視点撮影による大規模データセットを作成し、複数品目に対して品質分類モデルを構築・評価する。
- 第5章：自動AI選果システム実現に向けた現在の課題と、導入事例のインタビューを通じた要件分析、および未来の研究方向性を示す。
- 第6章：本研究の結論をまとめ、今後の展望を示す。





図 1.5: 等級選別を行っている様子 (ヤマイシ果樹園にて撮影).



## 第 2 章

# 関連研究

農作物の自動選別システムに関しては、産業界において実用化されている事例が複数存在する。たとえば、システック株式会社のカメラ式選果機「多選くん」<sup>\*1</sup>は、果実の大きさや形、色をルールベースで判定し、任意の排出口に振り分ける仕組みを提供している。また、株式会社ニレコの「AP800 シリーズ」<sup>\*2</sup>では、あらかじめプログラムされた判断基準に基づき、高級青果物を含むさまざまな作物（ナス、キュウリ、トマト、長芋、ジャガイモなど）を外観検査できる機能を備えている。さらに、AI 技術を活用した市販システムも登場しつつあり、静岡県の小池農園では 3 万 6 千枚のキュウリ画像を学習した自動選別 AI を用いて 8 割程度の精度を達成している<sup>\*3</sup>。JA みっかびでは、みかん選別のための柑橘選果場を整備し、総額 80 億円をかけて構築した設備に内部品質センサーや腐敗果センサー、外部品質カメラを導入して数万個のみかんのデータを収集し、AI による等級分類を行っている（図 2.1）<sup>\*4</sup>。

学術研究の分野でも、農作物の鮮度評価や等級づけを行うための AI 技術が活発に検討されている。Rybacki らは量み込みニューラルネットワーク（CNN [4]）を用いてナツメヤシ果実の色や幾何学的特徴を分類基準とするモデルを開発し、93% の精度を達成した [5]。Fu らは、ResNet [6] や VGG [7] など複数の CNN モデルを組み合わせた多クラス果実分類手法を提案し、事前に YOLO [8] で果実を検出してから鮮度スコアを算出するパイプラインを構築している（図 2.2）[9]。Zheng らはマンゴーの等級づけに特化した CNN システムを開発し、高精度化のみならず、実行速度や学習データ量の最適化にも成功している [10]。

農作物の異常検出や欠陥検出についても研究が盛んである。Arango らや Wu らは、それぞれりんごに生じる傷や欠陥の検出を CNN で行い、90% 以上の認識精度を記録している [11, 12]。Yang らは ResNet [6] と YOLO [8] を組み合わせることでジャガイモ表面に生じる複数の欠陥を同時に検出する手法を提案し [13]、Zhou らはにんじんの表面欠陥検出において知識蒸留を取り入れ、軽量かつ高速なネットワークを実装している [14]。また、Mendoza らは病気や雑草、害虫な

---

<sup>\*1</sup> <https://premium.ipros.jp/systecserver/product/category/42215/>

<sup>\*2</sup> <https://www.nireco.jp/product/agricultural/external-sensor/ap800.html>

<sup>\*3</sup> <https://smartagri-jp.com/smartagri/1579>

<sup>\*4</sup> <https://mikkabimikan.jp/blog/60>





図 2.1: JA みっかびの大型みかん選果場.



図 2.2: [9] で提案された階層型深層学習モデルのパイプライン.

ど多様な異常を農作物画像から判定するアプローチ [15], Liu らは自己教師あり学習を用いた異常検知手法を果実に適用する研究 [16] を行うなど, 新たな方法論の開拓が続いている.

一方, これらの先行研究や既存製品の多くは単一視点または限定的な角度で作物を撮影した画像を主に扱っており, 形状や傷, 色むらなどの外観特徴を多視点から統合して把握する試みは限られている. とくに, 多視点かつ大規模に撮影された公開データセットは少なく, 研究ごとに撮影条件やデータ規模, ラベリング基準がばらばらであるため, 異なる研究成果を横断的に比較したり, 手法を流用・改良したりするのが困難である. さらに, 3次元点群を用いて作物の立体形状や表面の歪みまで考慮する研究はごく限定的であり, LIDAR や 3D スキャナーのコスト・運用負荷から実地導入は容易ではないという課題もある.



以上の背景を踏まえ、本研究では下記の 2 点を主な目的とする。

### 1. 3 次元点群情報の活用

作物の形状や体積、表面の局所的な歪みなどをより正確に把握するため、2 次元画像だけでなく 3 次元点群データも統合し、深層学習モデルの可能性を探る。点群の取得コストや前処理負荷と精度向上のバランスを検証し、実用化へ向けた課題を洗い出す。

### 2. 多視点データセットの構築

単一視点では見落とされがちな形状の特徴や傷をより精密に捉えるため、大規模に多方向からの撮影を行う。さらに、統一された撮影条件とラベリング基準で複数品目を撮影し、異なる作物や規格間でのモデル性能比較が可能なデータセットを整備する。

本研究によって構築される大規模・多視点データセットと、2D + 3D を活用した品質分類モデルの検証は、多様な研究成果を比較・流用しにくい現状の課題に対する一つの解決策を提示するものである。複数の作物を同一の撮影環境下で大量に収集し、サイズや等級など外観品質に基づくラベルを付与することで、従来より統一的かつ実用指向の議論が可能になると期待される。



## 第 3 章

# 3 次元点群情報を考慮した農作物の外観品質選別

### 概要

日本では、若年層の人口減少と高齢化に伴い、農業における労働力不足が深刻化している。十数件の農家訪問を通じて、農作物の選別作業が国内農業サイクルにおける省力化のボトルネックであることが確認された。特に、品質選別工程の自動化は十分に進展しておらず、依然として多くの場合、人手によって行われている。既存研究では主に画像分類による品質選別が試みられているが、実際の選別では表面の傷や色だけでなく、作物の立体形状も重要な要素であるため、画像のみを用いたモデルでは人手と同等の精度を達成することは困難である。本研究では、2 次元画像と 3 次元点群データを統合して入力とする選別 AI モデルを提案した。みかん農家の協力を得て、200 個のみかんを撮影・3D スキャンし、データセットを作成した。モデルには ResNet と PointNet を組み合わせ、画像のみ、点群のみ、両者を組み合わせたモデルの性能を評価した結果、画像と点群の双方を入力とするモデルが最も高い精度を示し、95% の精度での選別が可能であった。これにより、3 次元情報が選別精度の向上に寄与することが確認された。

### 3.1 本研究の背景

1 にて論じた通り、農業プロセスの省人化や機械化は喫緊の課題であり、その中でも選別工程が大きなボトルネックとなっている。近年は画像認識や画像分類技術を応用した農作物の自動選別機が一部の施設で導入されているが、国内ではまだ普及が進んでいない。従来の選別装置や AI システムは二次元画像を主な入力としており、平面的な情報をもとに等級分類を行っている。しかし、実際の選別過程では、作物の表面の傷や色だけではなく、立体的な形状が判断材料となっている。例えば、図 3.1 はヘタの部分がすぼんでいるみかんである。みかん表面の傷や色は問題ないが、歪んだ形のため規格外品と分類される。作物を 2 次元的にとらえる既存研究や製品 (特に単一視点からの撮影) では、作物の立体形状をうまく捉えられず、選別精度が落ちてしまう可能性が





図 3.1: 形が歪んでいるみかんの例.

ある.

そこで本研究では、従来の二次元画像に加え、三次元点群を併用して農作物を認識・分類する枠組みを提案する．三次元情報を活用することで、作物の形状や体積、局所的な歪みなどをより正確に把握でき、二次元画像のみでは見分けが難しかった外観差異を捉えられる可能性がある．具体的には、カメラ画像から得られる視覚的特徴に加え、三次元スキャナーなどで計測した点群データを入力とする深層学習モデルを構築する．三次元情報を活用することで、農作物の形状や体積などの空間的な情報を詳細に取得でき、これまでの二次元画像では達成し難かった高精度な選別が可能となる．加えて、より少ないデータ量で高精度な選別を実現することが期待される．

点群データとは、3次元空間における位置情報を持つ点の集合であり、実物の形状を数値化するためのデータ構造である．各点はX、Y、Zの3次元座標で定義され、色や光の強度などの追加的な物理的特徴を含むことがある(図 3.2 参照)．点群データは、自動運転車の環境認識、ロボティクスのナビゲーション、地理情報の精密測量、都市開発計画、文化遺産のデジタルアーカイブなど、多岐にわたる分野での応用が進んでいる．近年、iPhoneなどの一般消費者向け製品にも搭載



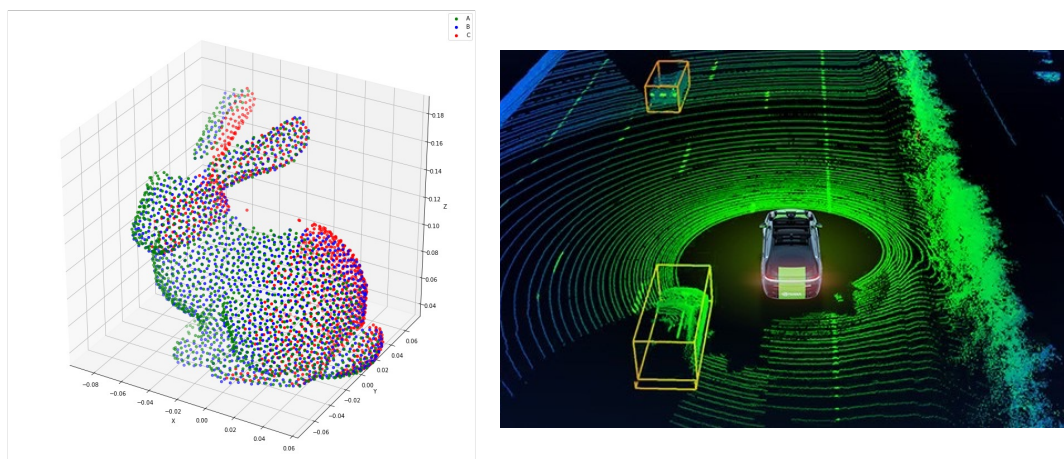


図 3.2: 左：実世界の物体から取得した構造化空間データを表す 3 次元点群の例\*<sup>1</sup>。右：自律走行車の LIDAR システムによる点群計測による周囲の車と環境の描写\*<sup>2</sup>。

されるなど、点群は広く普及しつつある。

本研究では、こうした点群活用の有効性を検証するため、画像と点群を同時に取得した農作物データセットを新たに作成し、深層画像処理モデル (ResNet) と深層点群処理モデル (PointNet) を組み合わせた選別 AI モデルを構築する。具体的には、みかんを対象として A 品・B 品などの等級分類を行い、画像のみ、点群のみ、両者を併用したモデルそれぞれの性能を比較する。これにより、三次元情報を加味することで高精度化やデータ量削減がどの程度期待できるのか、あるいは点群導入に伴うコスト・運用面の課題がどの程度生じるのかを明らかにする。

## 3.2 関連研究

点群データは以下に示す特性を持ち、その特性が点群データ処理の難易度を上げている。

- **不規則性**：異なるセンサーやスキャンの条件によって点の密度や分布は大きく異なり、ルールに基づいて固定された形状やサイズに整形されているとは限らない。
- **データ量の多様性**：点群データはしばしば大規模であり、非常に多くの点 (数百万から数十億点) を含むことがある。一方で、数百から数千しかない場合もある。また、各点が位置だけでなく、色や反射強度、時間経過など、多様な情報を保持することがある。
- **順序非依存性**：一般的に、点群データは順序に依存しない。点群データ内の点の順序を入れ替えても、点群の表すデータは同じである。
- **近傍の非決定性**：点群データは画像のような規則性のあるデータとは異なり、近傍が非自明

\*<sup>1</sup> <https://laempe.github.io/pyoints/tutorials/icp.html>

\*<sup>2</sup> <https://blogs.nvidia.co.jp/2021/03/04/lidar-sensor-nvidia-drive/>



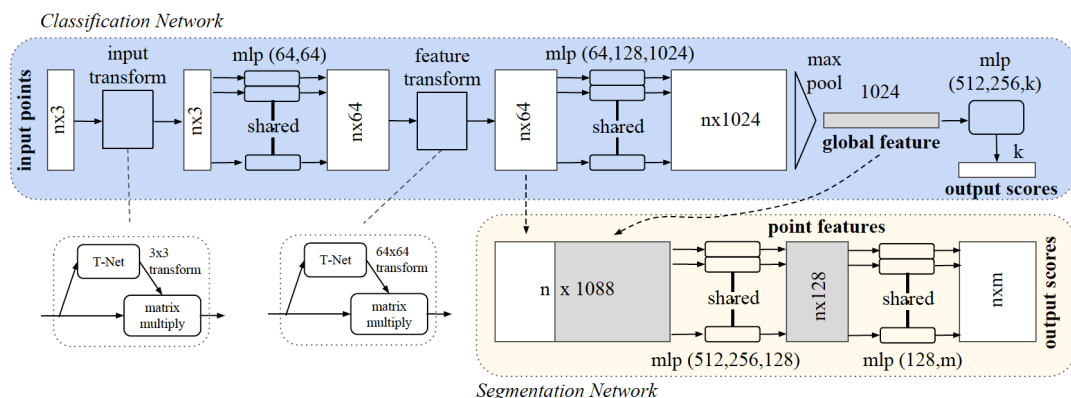


図 3.3: 入力、特徴変換、最大プーリング層を含む分類ネットワークと、共有 MLP と各点の出力スコアによるオブジェクト部分のセグメンテーションを行うセグメンテーションネットワークから構成される PointNet アーキテクチャの図 [18].

である。

これらの特性により、2次元画像のような規則的なピクセル配列とは対照的に、点群データの解析や処理は複雑かつ難しい。このような点群データならではの特性は、他の画像処理技術の直接的な応用を困難にしており、点群処理に特化した新たなアプローチが求められている。この点において、深層学習技術が注目されており、点群データ処理・分析に対して多くの深層学習手法が提案されている [17]。深層学習に基づく点群処理の最大の進歩は、データからの自動的な特徴抽出である。ニューラルネットワークは、点群の各点の局所的な構造からグローバルな形状に至るまで、あらゆる特徴を自動的に学習できる。これにより、分類、物体認識、セグメンテーションといったタスクが、従来手法よりも精度が高く、迅速に行えるようになった。また、深層学習は、ノイズや異常値に強い堅牢なモデルを構築することが可能である。従来手法ではノイズ除去や異常値の処理が大きな課題であったが、深層学習モデルはこれらの要素を自動的に認識・フィルタリングする能力を持つ。これにより、データの品質に依存することなく、一貫した性能を保持することができる。さらに、深層学習は不規則な点群データの扱いにおいても優れた柔軟性を発揮する。伝統的手法が特定の形状やモデルに依存していたのに対し、深層学習はデータ内の任意の構造を学習し、より多様なシーンやオブジェクトに対応可能である。例えば、自然環境や複雑な都市景観の点群データにおいても、高い精度で特徴を捉えることができる。このように、深層学習は点群データ処理の分野において、データの扱いやすさ、処理速度、汎用性において大きな進歩を遂げた。

PointNet [18] は、点群を入力とする代表的な深層分類・セグメンテーションモデルである。PointNet は、点群の順序非依存性を扱えるようにアーキテクチャが設計されている。図 3.3 に、PointNet のアーキテクチャ図を示す。PointNet の重要な特徴は、入力される点の順序がネットワークの出力に影響を与えないアーキテクチャを持つことである。また、PointNet は、点群データが平行移動や回転をしても、同じ出力結果を保証するため、ネットワーク内に変換層 (Transform



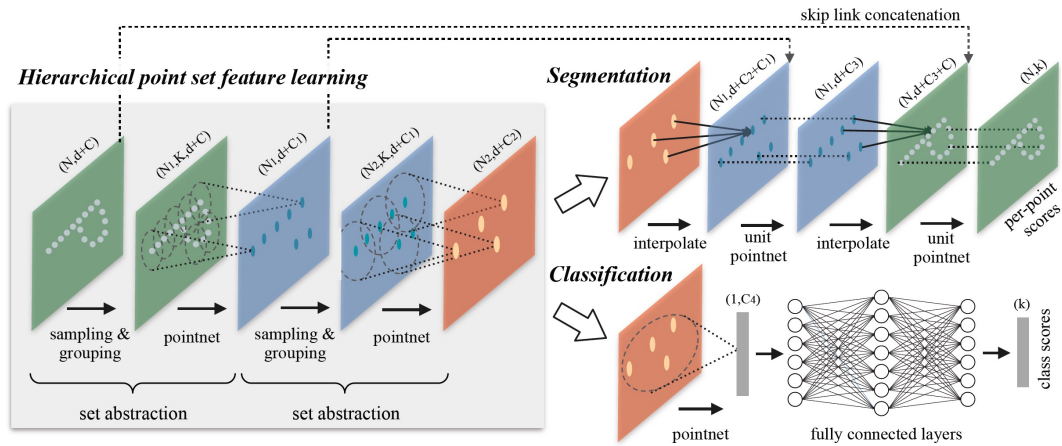


図 3.4: PointNet++ のアーキテクチャ図 [19]. 点集合特徴学習のための階層型ネットワークと、セグメンテーションおよび分類タスクへの応用を示す. 左側は階層的な特徴抽出プロセスを示しており、サンプリングとグルーピング操作に続いて、複数の PointNet 集合が抽象化される. 右側は、これらの学習された特徴量のセグメンテーションへの応用を、スキップリンク連結と補間ステップによって、また分類への応用を、ポイント毎のスコアとクラススコアを出力する一連の完全連結層を通して示している.

Layer) が導入されており、点群データをランダムな姿勢に変換することで、入力回転や平行移動に対するロバスト性を向上させている. 点群データはまず、共通の多層パーセプトロン (MLP) と変換層を通過し、それぞれの点に対する特徴量へと変換される. 続いて、これらの点ごとの特徴量に対して最大プーリング操作を行い、一つのグローバル特徴量を抽出する. このグローバル特徴量は、次に MLP によって分類が行われる. このプロセス全体を通じて、アーキテクチャは対称関数 (Symmetric Function) を形成し、入力される点群の順序に関わらず一貫した出力を生成するように設計されている. セグメンテーションタスクの場合、PointNet はそれぞれの点の特徴量とグローバル特徴量を統合し、混合された特徴量を用いて各点のカテゴリを識別する. グローバルな情報を考慮した上での点ごとの特徴量に基づくこのプロセスにより、全体を考慮したセグメンテーションが可能となる. PointNet のアーキテクチャには、点の局所的な特徴、つまり近接関係を考慮できないという欠点がある. これは、画像における MLP が位置関係を失うことに相当し、点群データの中での局所的な構造やパターンを捉えることができない. この問題は、後に PointNet++ などの進化したモデルで局所的な特徴量を抽出を可能にすることで解決された.

PointNet++ は、画像処理の領域で革新的な役割を果たした畳み込みニューラルネットワーク (CNN) [4] の概念を、点群データ処理へと応用したモデルである. CNN は画像認識や分類において、画像のピクセル間の空間的関連性を捉える能力により、目覚ましい精度向上をもたらした. その先駆者として知られる AlexNet [20] は、ImageNet データセット\*<sup>1</sup>に対する分類精度を大幅

\*<sup>1</sup> <https://www.image-net.org/>



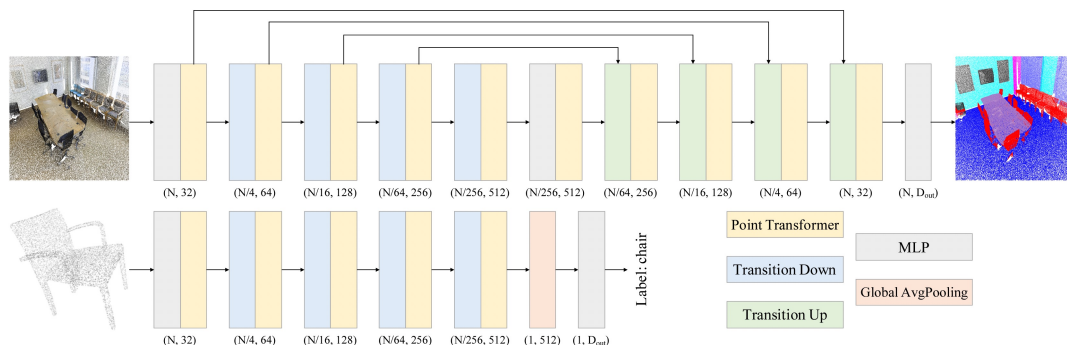


図 3.5: 3D 点群解析のための Point Transformer ネットワークの概略図 [21]. セグメンテーションと分類タスクのための 2 つの異なるアーキテクチャが説明される. 右上のセグメンテーションされた点群によって可視化されるように, 正確なセマンティックセグメンテーションに不可欠なマルチスケールのコンテキスト情報を効果的に捕捉するため, 接続をスキップする U-Net のような構造を利用する. 下部は分類アーキテクチャを示しており, 愚直なフィードフォワード構造を採用している.

に向上させた. 同様に, 点群データを扱う際にも, 類似したアプローチが必要とされている. 点群データは不規則な形で存在するため, 画像データのように均一なグリッド上に整列していないという特性がある. PointNet++ は, この問題に対応するため, 空間を一定の範囲の球体に分割し, それぞれの球体内で点の局所的な特徴を PointNet を用いて捉える. このアプローチにより, 点群データの中での点間の空間的な近接関係を把握し, それを特徴量として利用する. 図 3.4 に, PointNet++ のアーキテクチャを示す. 点群に対して畳み込みを適用する際に, 一つ問題となるのは, 点の密度が不均一であることである. 点の密度がまばらであることにより, 一貫性のある局所特徴量の抽出が困難になってしまう. PointNet++ では, この問題を解決するために, 局所的な点の数を一定に保ち, かつ局所特徴量を失わないためのサンプリングが施される. これにより, 点の密度変動に影響されることなく, 安定した特徴抽出を実現する. 次に, PointNet++ は空間内の一定の範囲を球で切り出し, その球の内部に含まれる点群に対して PointNet を適用する. これにより, その領域内の局所的な特徴を抽出している. サンプリングと局所球内に PointNet を適用する操作を複数回繰り返すことで, PointNet++ は全体の特徴量を抽出し, この特徴量を基盤として, その後の点群分類やセグメンテーションを行う. セグメンテーションを行う際は, PointNet++ は画像処理における U-Net のようなスキップ接続を用いて, セグメンテーションタスクにおける各点の特徴を得る. このスキップ接続により, 異なる解像度の特徴マップを組み合わせることで, 詳細な局所的情報を維持しながら, 全体のコンテキストを考慮した特徴表現を生成することができる.

Point Transformer は, Transformer [22] のアーキテクチャを点群データの処理に適用した手法である. Transformer は元は自然言語処理のために開発され, 単語の並びに基づいてテキストの意味を把握する際に非常に優れた性能を発揮した. その後, Vision Transformer (ViT) [23] が



登場し、画像処理においても同様のアプローチをとって高い性能を示した。画像のピクセルやパッチを単語のように扱い、単語間の関係性を捉えることで、従来の CNN を超える精度を達成した。Point Transformer は、このような Transformer の考え方を点群データに応用したものである。図 3.5 に、Point Transformer のアーキテクチャを示す。従来の位置符号化 (Positional Encoding) の代わりに、点の相対位置を変換して加算することで、点群の空間的な構造を捉える。点群内の各点がそれぞれに対話するように情報を交換しあい、最終的な出力に至るまでにグローバルな特徴を含んだ各点の特徴を獲得することができる。セグメンテーションタスクであれば U-Net [24] のようなスキップコネクションを用いた構造をしており、分類であれば単純なフィードフォワード構造をしている。Point Transformer は、セグメンテーションや分類といったタスクにおいて、当時の State of the Art を達成した。

PointMLP は、画像分類における ResNet [6] に相当する。ResNet は、残差接続 (Residual Connection) を使用することで、より深いネットワーク構造を実現し、それにより画像分類の分野において顕著な精度の向上を達成した。同様に、PointMLP も点群データの精度、推論速度、学習速度において優れた性能を発揮する。図 3.6 に、PointMLP のアーキテクチャを示す。PointMLP は、局所的な点群に対してアフィン変換を施し、これを深層の MLP とプーリングによって処理する。残差接続を活用することで、より深いネットワーク構造を可能にし、その結果、精度の高い学習が可能となる。これにより、PointMLP は 3D 空間内の複雑な形状や構造を持つ点群データの特徴を効果的に抽出し、分類やセグメンテーションといったタスクにおいて優れたパフォーマンスを示す。

本研究の目的は、PointNet を活用して、従来の二次元画像に基づく AI 選別機に比べて、点群データを直接処理することにより、農作物選別の精度を向上させることである。点群データから得られる三次元の空間情報を用いることで、形状や体積といった従来の画像では捉えにくい農作物の特性を詳細に分析する。この手法により、より少ないデータ量で高い精度の選別を実現し、農業現場での選別効率の向上とコスト削減を目指す。また、将来的には三次元データの特性を生かした新たな特徴抽出方法を開発し、農作物の微細な欠陥や品質の違いをより正確に識別できるようにする。

### 3.3 提案手法

本研究では、農作物の品質選別を自動化し、農業の人手不足問題に対処する選別システムを開発することを目指す。既存の AI 選別機は一般に画像データのみを利用する。本研究では、画像に加えて三次元点群データを含んだ作物データセットを作成する。また、画像と点群両方を入力として用いたマルチモーダルなモデルを構築し、精度検証を行う。

高知県岡宗農園様の協力のもと、人手で選別された 214 個のみかん「はるか」の画像と点群データを収集した (表 3.1)。これらのみかんは見た目の品質に基づいて A 品 (101 個)、B 品 (99 個)、C 品 (14 個) にあらかじめ選別され、ラベル付けされている (図 3.7)。画像データは各みかんを六つの異なる角度から iPhone カメラで撮影し、一個当たり六枚の画像を取得した (図 3.8)。点群



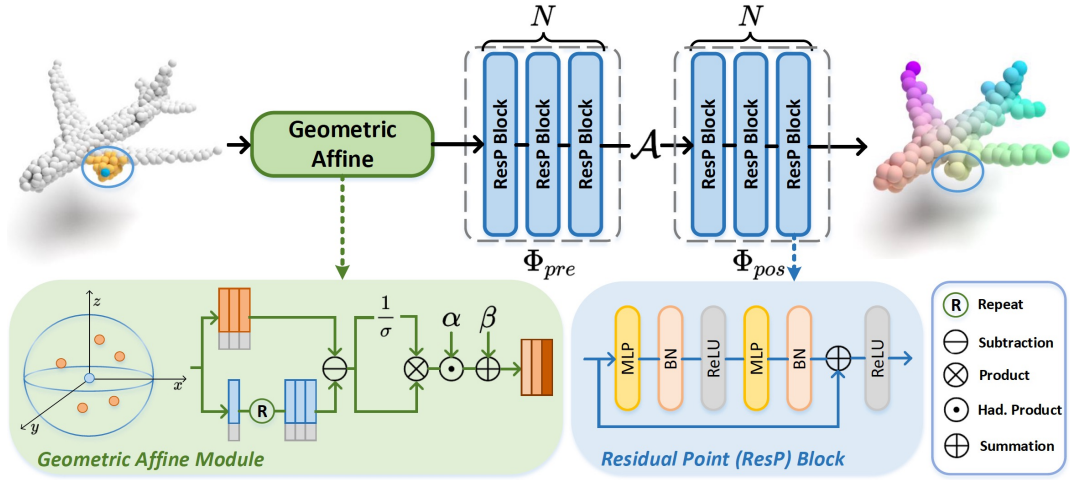


図 3.6: PointMLP フレームワークのアーキテクチャ [25]. 入力データの幾何学的アフィン変換から始まり、ポイントの特徴を徐々に洗練していく一連の Residual Point (ResP) ブロックが続く。左下が「Geometric Affine Module」で、アフィン変換の詳細を示している。中央の「ResP ブロック」構造は、多層パーセプトロン (MLP)、バッチ正規化 (BN)、整流線形ユニット (ReLU) を含むネットワーク層を示す。

データは、みかんの上部と下部をそれぞれ 3D スキャナーでカラーキャンし、みかん一個あたり二つの点群ファイルを取得した (図 3.9)。すべての画像データは背景の差異や撮影者の個人差を排除するために、背景を切り抜き、サイズを標準化した。

### 3.4 実験

本研究では、2 次元画像と 3 次元点群の両方を入力として、みかんの外観品質を分類するマルチモーダルモデルを構築した。畳み込みニューラルネットワークの一種である ResNet [6] と、点群データを処理するためのネットワークである PointNet [18] の二つを用いる。ResNet は、画像からの特徴抽出に優れており、PointNet は三次元空間データの局所および全体的な特徴を捉える能力がある。画像データを ResNet に入力し、点群データを PointNet に入力する二系統のネットワークを用意し、それぞれの特徴ベクトルを結合した後、最終的な全結合層に通して等級を推定する。これにより、画像のみ、点群のみの情報を別々に抽出するのではなく、2 つのモダリティから得られる特徴を同時に学習し、統合的に判定を行うことを目指す。

比較実験としては以下の 3 種類のモデルを用意し、それぞれの性能を比較する。

#### 1. 画像のみモデル (ResNet-18 単体)

ImageNet で事前学習済みの ResNet-18 を初期値として利用し、最終層のみ本データセットに合わせて再構築して学習する。



表 3.1: 構築されたデータセットの詳細.

等級	訓練データ数	テストデータ数
はるか A 品	81	20
はるか B 品	79	20
はるか C 品	14	-

## 2. 点群のみモデル (PointNet 単体)

点群処理を目的として提案された PointNet を用い, 3 次元形状や局所幾何情報を抽出し, 最終的にみかんの等級を分類する.

## 3. マルチモーダルモデル (ResNet + PointNet)

画像は ResNet-18 へ, 点群は PointNet へ入力し, それぞれの最終手前の特徴ベクトルを結合 (concat) して全結合層に入力する. 同一の損失関数 (クロスエントロピーロス) を通じて誤差逆伝播を行い, ResNet と PointNet の両方を End to End で同時学習する.

画像データは前処理として正規化を行い, 6 枚の撮影画像を平均合成したものを 1 サンプルとした. C 品の数が少ないため, 本研究でのモデルの訓練と評価には A 品と B 品のみ合計 200 個のみかんのデータを使用した. これらを 128 個の訓練データ, 32 個の検証データ, 40 個のテストデータに分割し, それぞれのデータセットでモデルの訓練, 検証, 評価を行った. また, データの均一な分布を確保するため, 各カテゴリーからランダムにサンプリングし, モデルの汎用性とロバスト性を検証した. 点群データとしては, 各みかんの点群から 4096 個の点をランダムに抽出し, 上部と下部の点群を結合したものを PointNet の入力とした. 各点は  $(x,y,z,r,g,b)$  の 6 次元の情報を用いた. 学習の際, 学習率は 0.0001, バッチサイズは 128, 損失関数はクロスエントロピーロス, 最適化手法は Adam を用いた.

本実験では, 評価指標として Accuracy (精度) とマッシュューズ相関係数 (MCC) を採用した. Accuracy はモデルが全データに対してどれだけ正確に分類を行えたかを示す指標であり, 一般的にはモデルのパフォーマンスを評価する際に広く使用されている. 一方, MCC はバイナリおよびマルチクラス分類問題においてクラス間のバランスを考慮した指標であり, データセットが不均衡な場合の評価に特に有効である [26]. MCC は -1 から 1 までの値をとり, 1 は完全な予測, 0 はランダム予測, -1 は完全に逆の予測を意味し, この指標により, モデルの予測がどれだけ実際のラベルと相関しているかをより正確に評価することができる. これらの指標を用いることで, モデルの性能を総合的かつバランス良く評価することが可能となる.

以上のように, 画像のみ・点群のみ・画像+点群 (マルチモーダル) の 3 通りのモデルを同一条件で学習・評価し, 三次元情報の追加が外観品質選別の精度向上に寄与するかを明らかにする. 特にマルチモーダルモデルでは, 2 つのネットワークを同時に学習させることで, 画像と点群の潜在的な相補関係を最大限に引き出すことを狙いとしている.



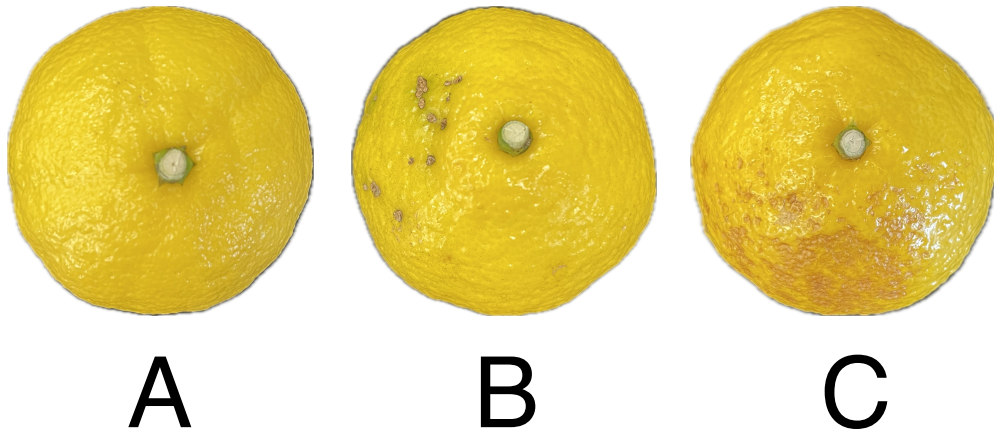


図 3.7: それぞれみかん A 品, B 品, C 品の例. 背景はトリミングされ, サイズは均一である.

### 3.5 結果と考察

表 3.2 に示すとおり, 今回の実験では, 画像と点群を組み合わせたマルチモーダルモデルが最も高い分類性能を示した. 具体的には, Accuracy が 0.950, MCC が 0.905 という数値を達成し, いずれも他の単一モーダルモデルを大きく上回っている. 一方で, 画像のみを入力とした ResNet-18 単体モデルでは精度 0.875, MCC 0.759 と, 単体としては十分に高いパフォーマンスを示したものの, マルチモーダルモデルほどには及ばなかった. また, 点群のみを入力とする PointNet 単体モデルは精度が 0.650, MCC が 0.306 にとどまり, 形状情報だけでは等級を精確に区別することが難しいことがわかった. さらに, 人間 (著者) による目視判定の精度は 0.800 (MCC 0.612) であり, マルチモーダルモデルの 0.950 という値が, 熟練者の判断を超える可能性を示唆する結果となっている.

これらの結果からは, まず, 三次元点群と二次元画像の双方を同時に活用するマルチモーダル学習が, 外観品質選別タスクにおいて非常に有効であることがうかがえる. 画像だけでは見分けが難しい微妙なへこみや変形を, 点群の立体形状や体積情報によって補完できる一方, 点群単体では捉えきれない色やテクスチャに関する情報を画像が補うことで, 総合的な外観評価の精度が向上していると考えられる. このように, 二つのモダリティが相互に特徴を補完し合い, 単一モーダルでは達成できない識別能力を発揮していることが, マルチモーダルモデルの大幅な精度向上につながったと推察される.

点群のみを用いたモデルの精度が伸び悩んだ点については, スキャン時のノイズや, みかん表面の形状が品種や個体差によって微妙に異なることが影響している可能性がある. PointNet は 3 次元点群を処理する代表的な手法であるが, 局所的な幾何情報をより巧みに抽出する PointNet++



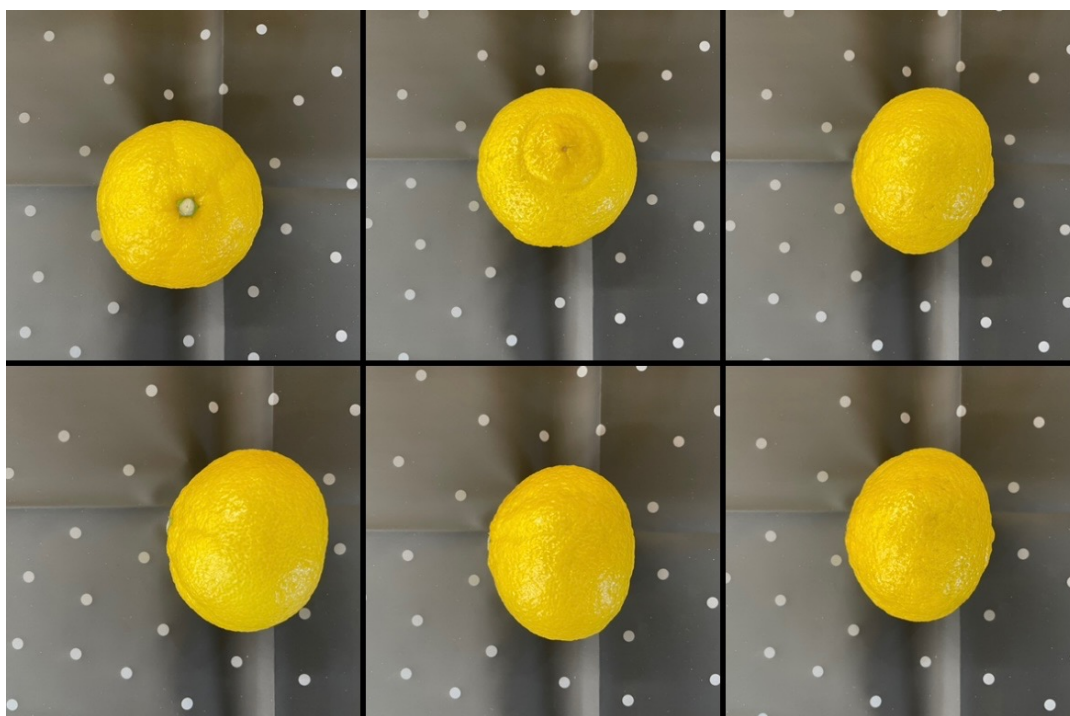


図 3.8: 6 方向から撮影したみかんの画像.

や、Transformer 型のモデルを利用すれば、形状由来の特徴をさらに精密に取り出せる可能性がある。また、計測に用いる 3D スキャナーの解像度や配置にも改良の余地があり、より高精細に作物表面の凹凸や傷を取得できれば、点群単体でも高精度化が見込める可能性がある。

一方、画像モデル単体であっても、今回の研究では 0.875 という高い精度を示したことから、カラー画像による深層学習が外観評価に対して有効であることも再確認できる。多くの農業現場で比較的手軽に導入可能なカメラ技術や CNN ベースのアルゴリズムだけで、一定水準の自動選別を行える見通しが立つため、コスト面や設備面を考慮すれば、まずは画像単体のモデルを導入し、必要に応じて点群を追加する段階的アプローチも実用的な選択肢になり得る。

加えて、本研究では人手による判定も行ったが、著者個人であっても精度が 0.800 に達し、ベテラン生産者であればさらに上回る可能性がある。現場における選別基準は多岐にわたり、熟練者ごとに差異があるため、AI モデルが実運用に組み込まれる際には、なぜその判定に至ったのかを可視化・説明できる枠組みや、判定結果を最終的に人間が承認できるフロー作りが重要になる。マルチモーダルモデルが高精度だからといって、すぐに現場へ導入できるとは限らないが、少なくとも今回の結果は、三次元点群と二次元画像を併用する手法に実用上の大きな可能性があることを示唆している。

以上の考察から、三次元情報を活用することによって画像モデルを大きく上回る精度を獲得できる一方で、三次元計測のコストやノイズ対策、現場作業との整合性といった課題も依然として残



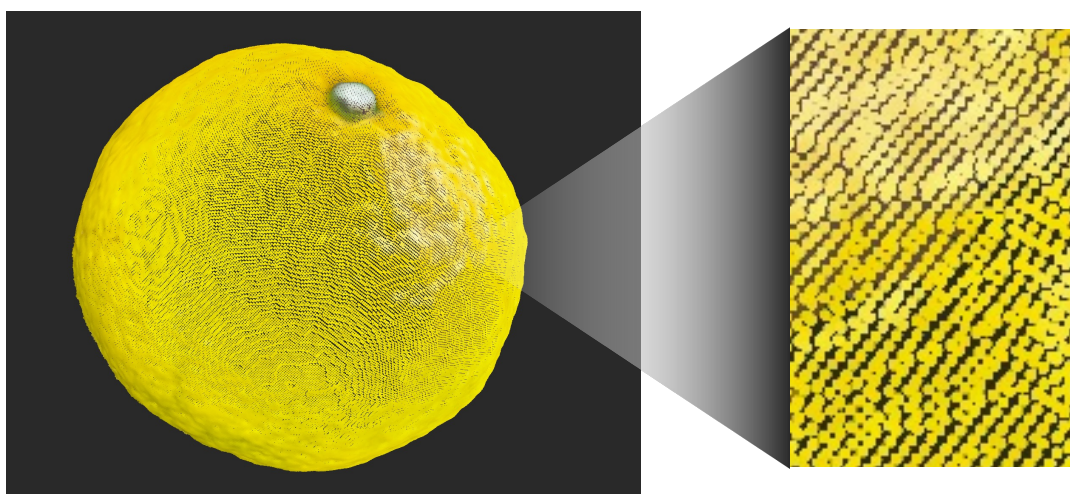


図 3.9: 取得したみかんの点群データの例.

表 3.2: 各モデルの定量的評価

モデル	Accuracy (↑)	MCC (↑)
画像のみ (ResNet-18)	0.875	0.759
点群のみ (PointNet)	0.650	0.306
画像 + 点群	<b>0.95</b>	<b>0.905</b>
人 (著者)	0.800	0.612

されていることがわかった。これらの点を踏まえ、本研究ではさらに多様な品目や撮影環境を想定した実験や、PointNet++ などの改良モデルへの適用も検討する必要がある。

### 3.6 未来の研究方向性

研究では、二次元画像と三次元点群をそれぞれ用いた農作物選別モデルを比較した結果、現時点では画像ベースのモデルのほうが高い精度と導入のしやすさを両立できることが明らかになった。一方で、点群データは農作物の形状や体積、微細な凹凸を直接的に反映する特性を持ち、二次元画像のみでは把握しきれない外観上の差異を捉える可能性がある。そのため、三次元情報の活用は外観品質選別の高精度化に寄与する有望な手段であるものの、スキャナーなどの取得デバイスやアルゴリズムへの計算リソースが高額に及ぶという課題が依然として大きい。特に、小規模農家や限られた設備投資しか行えない生産者にとっては、ハードウェア調達やデータ処理フローの複雑化が導入の障壁となる。

しかし、近年は 3D センサーの低価格化やコンパクト化が進み、消費者向けデバイスにも点群計



測機能が搭載され始めている。アルゴリズム面では、PointNet++ [19] や PointMLP [25] など、より洗練された局所特徴の抽出が可能なモデルが提案されており、高精細な点群データが得られれば、単純な形状の相違だけでなく、表面の微細な傷や凹凸も捉えやすくなるだろう。ハードウェアとソフトウェアの両面で性能とコストの最適化が進めば、三次元点群の導入ハードルは段階的に下がり、より多くの生産現場で活用が検討される可能性が高い。

今後の研究としては、第一に、多様な作物や品種を対象にして三次元情報がどの程度有効に働くかを検証する必要がある。形状や表面の特徴が作物によって大きく異なるため、汎用的に適用できるモデルの構築が重要になる。第二に、高性能な点群処理モデルの開発だけでなく、安価なセンサーや簡便な測定手法と組み合わせることで、導入コストを抑えながら高精度な外観評価を実現する方策を探ることが求められる。第三に、二次元画像と三次元点群をさらに密接に統合するマルチモーダル手法の洗練が挙げられる。色や模様などの画像特徴と、立体的な歪みや大きさを示す点群特徴を有機的に融合することによって、より柔軟な選別基準にも対応できるモデルを目指すことができる。

このように、三次元情報は本研究で示した課題を克服すれば、農作物の外観品質評価を飛躍的に高精度化する可能性を秘めている。アルゴリズム、ハードウェア、実装コストのいずれにおいても技術革新が進みつつあるため、今後はこれらを総合的に活用し、小規模農家でも無理なく利用できる三次元点群活用の選別システムを開発することが期待される。こうした取り組みが進めば、最終的には二次元画像ベースの選別機と三次元点群活用のハイブリッド型が標準的なスタイルとなり、より多くの現場で選別工程の省力化と精度向上が実現されるだろう。

### 3.7 結論

国内の農業分野では若年層人口の減少と高齢化が進行しており、その結果として深刻化する人手不足の中で、特に作物の選別工程が生産現場における省人化のボトルネックとして浮上している。年間約 3300 億円ものコストが選別作業に投じられているにもかかわらず、品質選別工程の機械化や自動化はほとんど進んでおらず、高額な設備投資や大量データの学習コストが普及の障壁となっている状況がある。本研究では、既存の二次元画像に加えて三次元点群を取り入れることで、農作物の形状や表面の歪みまで詳細に捉えられるモデルを構築し、高精度な外観品質選別を実現する可能性を探った。具体的には、みかんを対象に ResNet を用いた画像モデル、PointNet を用いた点群モデル、そして両者を同時に学習するマルチモーダルモデルを比較した。その結果、画像のみのモデルが Accuracy 0.875、点群のみのモデルが 0.650 にとどまったのに対し、両モダリティを併用するマルチモーダルモデルは Accuracy 0.950、MCC 0.905 と大幅に精度が向上した。これは、表面の色や微妙な傷を捉える画像の強みと、立体的な形状や体積情報を反映する点群の強みが互いを補完し合い、単一モーダルでは得られなかった識別能力を引き出していることを示唆している。

一方で、点群導入には 3D スキャナーや計測装置のコストがかかるほか、ノイズや欠損の処理といった課題が残るため、すべての生産現場で三次元情報の導入が即座に適合するわけではない。し



かし、近年の技術進歩によって、3D センサーや深層学習アルゴリズムは低価格化・高性能化が進んでいることから、設備投資の負担軽減や精度向上が今後さらに期待できる。特に、PointNet++ や PointMLP などの高度な点群処理モデルを活用すれば、単純な形状だけでなく表面の細かな凹凸や傷も詳細に把握し、より多様な作物や品質基準に柔軟に対応できる可能性が高い。

実地導入を見据えると、小規模農家でも手が届く低コスト機器の開発や、異なる品種や産地の基準を扱うための柔軟な学習・転移学習スキームが重要になる。多視点画像と点群を効率的に取得し、高精度モデルを簡易に再学習できる基盤モデルを構築することで、農家それぞれの選別基準や多品目生産にも対応できる。こうした要件を満たすシステムが実現すれば、二次元画像の気軽さと三次元点群の高精度を兼ね備えた選別装置を比較的安価に導入できるようになり、国内農業の省力化と効率化に大きく貢献するだろう。本研究で示したような三次元情報とのマルチモーダル活用は、外観品質選別の高度化に非常に有効なアプローチであることが示唆された。同時に、点群取得や運用コスト、現場との調整など課題は多いが、技術進歩と実用面の連携を深めることで、近い将来にはより高精度かつ導入しやすい AI 選別システムが普及していくと考えられる。



## 第 4 章

# 農作物選別作業の自動化に向けたデータセット構築と品質分類モデルの開発

### 概要

日本国内の農業では若年層人口の減少と高齢化が進んでおり、特に農作物の選別工程における労働力不足が深刻化している。十数回の現地インタビューを通じて、品質選別が省力化の大きなボトルネックであることが明らかとなった。本研究ではこの課題を解決すべく、複数方向から作物を撮影できる専用の撮影機械を新たに開発し、レモン・みかん・じゃがいもを対象に数千枚規模の画像データを収集した。取得した大規模な多視点データセットを用いて深層学習モデル（ResNet-18）を訓練し、3等級に基づく分類タスクを実施した結果、レモンで 85.3%、みかんで 89.7%、じゃがいもで 84.3% という高い精度を達成している。これにより、多視点情報を活かした深層分類モデルが農作物選別の自動化に有用であることが示唆された。今後は本手法をさらに拡張し、実地導入に向けてコスト面や説明可能性、異常検知機能などを含む総合的なシステムを構築することで、農業現場の労働負担を軽減し、国内農業の生産性向上に寄与できると期待される。

### 4.1 本研究の背景

1 にて論じたように、国内農業は人手不足が深刻であり、特に作物の選別工程が省人化のボトルネックとなっている。そのため、画像処理技術などを活用した選別作業の自動化が急務である。このような選別自動化の遅れの一因として、大規模かつ高品質な農作物画像データセットの不足が挙げられる。農作物の等級評価は多様な基準に基づくため、単一視点の静的な画像だけでは把握しきれない情報が多く、色合いや傷の有無に加えて、形状や表面の微妙な凹凸など多方向からの情報収集が極めて重要である。にもかかわらず、既存の公開データセットは大半が単視点または限られた撮影条件下で得られたものであり、多視点情報を体系的に取得した大規模データが不足している。そのため、研究ごとに撮影環境や作物の品種が異なり、学習したモデルやその評価結果を横断的に比較・検証したり、新たな応用へ容易に転移したりするのが難しい状況となっている。



そこで本研究では、こうした課題を克服し、農作物の等級選別に特化した多視点データセットを新たに構築する。具体的には、レモン、みかん、じゃがいもという複数の品目を同一の撮影環境下で大量に撮影し、サイズや外観基準に基づいて等級ラベルを付与したデータセットを整備するとともに、このデータを用いて高精度な品質分類モデルを開発・評価する。多視点情報を取り入れることで、単一視点では捉えきれない傷や形状の差異を補足し、深層学習モデルの分類性能を向上させることを目指す。また、複数品目を同時にカバーすることで、異なる作物間でのモデル精度を比較したり、汎用的な基盤モデルを構築したりする足がかりとする。最終的には、このデータセットとモデルを活用することで、農作物の等級選別工程を高精度かつ自動化できるシステムの実装と普及に貢献し、農業の省力化と効率化を促進することを狙いとしている。

## 4.2 関連研究

農作物の自動選別システムに関する研究や製品開発は近年活発化しており、畳み込みニューラルネットワーク (CNN) を用いた画像分類や欠陥検出、YOLO [8] などの物体検出手法を組み合わせた鮮度判定など、多彩なアプローチが提案されている。実際に高精度なモデルが報告される一方、多くの既存研究や商用システムでは単一視点の画像を主に使用し、限定的な撮影条件や品目に特化しているのが現状である。そのため、形状や傷、色むらなどの外観特徴をより正確に捉えるために有効とされる多視点からの大規模データは依然として不足している。さらに、研究ごとに撮影環境やデータ規模、ラベリング基準が大きく異なるため、開発されたモデルを他作物や異なる条件へ横断的に適用するには困難が伴う。

従来研究や既存製品では、単一視点あるいは限定的な角度で撮影された作物画像を対象にした品質評価が主流となっている。しかしながら、形状や傷、色むらなどの外観特徴をより正確に把握するためには、多視点からの撮影によるデータが有効であるにもかかわらず、多視点で大規模に撮影された農作物データセットはほとんど公開されていない。また、研究ごとに撮影条件やデータ規模、ラベリング基準が大きく異なるため、異なる研究結果を横断的に比較したり、手法を流用・改良したりするのが困難な状況にある。そこで本研究では、多視点からの農作物画像を大規模に収集し、サイズや等級に基づく品質分類ラベルを付与したデータセットを構築することで、これらの課題解決を目指した。撮影の段階から高精度なカメラや照明環境を整え、同一条件下でレモン・みかん・じゃがいもを大量に撮影しており、データセットの有用性を示すために分類モデルを学習させ、精度検証を行った。

## 4.3 データセット作成

### 4.3.1 撮影機械の設計と開発

本研究では、農作物を多視点から効率的に撮影するための専用機械を設計・開発した。従来の農産物撮影では、手動で作物を回転させる、あるいは各方向から個別にシャッターを切る必要があったが、これらの作業は撮影時間の増大と作業者の負担増につながっていた。そこで、本研究では以



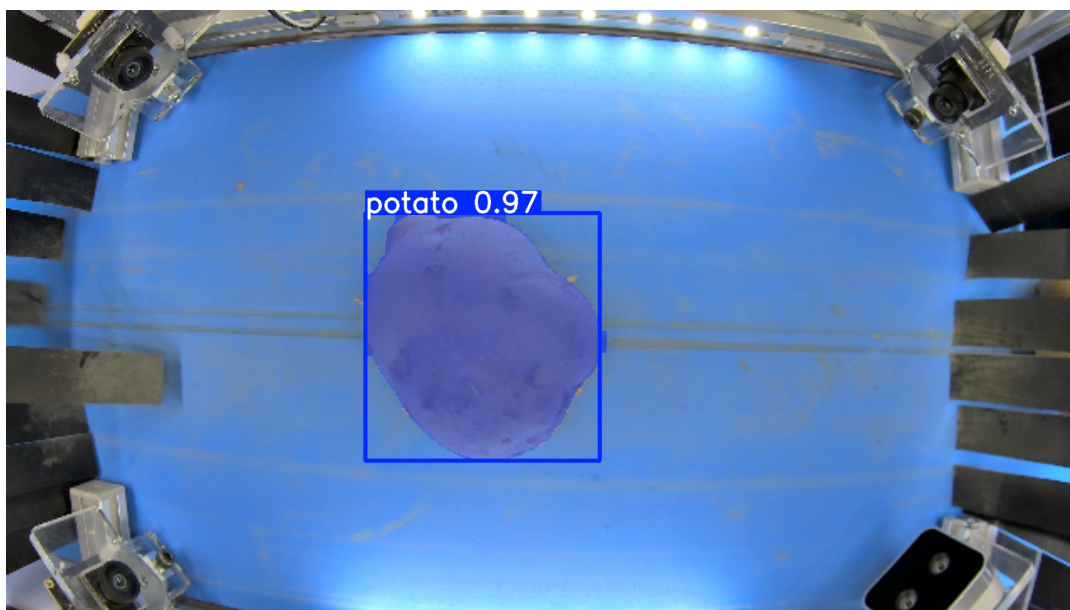


図 4.1: 作物検出の例 (じゃがいも).

下の設計方針に基づき撮影機械を開発した.

本研究で開発した撮影機械の設計では、まず多方向撮影の効率化を念頭に置き、作物を一定速度で搬送するコンベアに合わせて複数のカメラを金属製の箱状装置内に配置する構造を採用した. 撮影時に作物を回転させることなく、一度のシャッター操作で複数方向からの画像を同時に取得できるようになっている. さらに、外光を遮断する金属製の箱の内部には均一に LED 照明を配置し、環境光の変動や作物ごとの露出差を最小限に抑えることで、撮影条件を一定に保つ工夫を施した. また、撮影に使用するカメラとして HDR 対応の機器を選定し、高コントラストな場面でも階調情報をできるだけ保ちつつ、LED 照明からの反射やハレーションを抑制している.

撮影後の画像は、カメラ画像から作物を認識し、作物部分のみを切り出した上でデータセット化する. レモン、みかんを撮影した際には、背景差分法を用いて作物を自動的に抽出したが、背景との差が小さい個体やカメラアングルの変化によって検出精度が低下するケースが散見された. そこでじゃがいもでは、より高精度な検出を実現するために物体検出モデルである YOLO [8] を用いてバウンディングボックスを生成し、その結果をもとに画像を切り出す方式に切り替えた. 図 4.1 には、YOLO を用いてじゃがいもを検出し、バウンディングボックスで囲んだ実例を示している.

図 4.2 に本撮影機械を示す. 金属製の箱は撮影エリアと制御エリアが分離された 2 層構造になっており、下部にカメラ群 (最大 5 台まで搭載可能)、上部に制御用 PC や照明用電源などを格納している. 機械の制御は全てタッチパネルで行えるようになっている. また、コンベア部分にはミスミ社製のベルトコンベアユニットを用い、搬送速度は自由に変更できる. 撮影時は、1 時間あたり 1500 個程度の農産物を撮影できるように設定されており、大量の画像を効率的に取得する





図 4.2: 開発した農作物撮影機械.

ことが可能となった.

#### 4.3.2 データ収集の協力体制

本研究における大規模なデータ収集は、広島県庁農林水産局、JA 広島果実連、および東広島市安芸津町にてじゃがいもを栽培する中市後久夫氏の全面的な協力を得て実施された。まず、農家や選果場の担当者との連携を図り、作物の品種や収穫・選果の時期を調整してもらうことで、十分なサンプル数を確保できるようスケジューリングを行った。

データ収集の際には、各施設に撮影機械を持ち込み、選果作業の合間を縫って撮影を行った。各





図 4.3: みかんのデータ取得の様子 (JA ひろしま くれ選果場にて)。機械の奥で選果作業が行われている。

農産物のデータ取得にあたっては、まずレモンとみかんはサイズ選果機によって仕分けを行い、じゃがいもについては重量を測定してサイズ別のサンプルを確保した。次に、熟練した農家の目視による等級選別 (A 品, B 品, C 品など) も同時に実施し、傷や変色、形状の乱れなどの観点から等級付けを行った。このように、機械的なサイズ分別と人の判断による等級選別という二段階の仕分けを組み合わせることで、より価値の高いデータセットの構築を目指した。

図 4.3 と図 4.4 は、実際にデータを取得している様子を示している (撮影時の風景写真)。撮影時は、コンベアの入り口側で作物を順番に流し込む担当者と、出口側で撮影後の作物を受け取る担当者の二人体制で行った。作業者がコンベアに載せ、撮影機械の中を通過する間に多方向から画像を取得する。その後、もう一人の作業者が出口側で作物を受け取り、次々と新たな作物を流し込むことで、スムーズな撮影サイクルを維持することができ、大量のデータを短時間で取得することが可能となっている。なお、本研究での撮影機械の作成、データセットの収集・ソフトウェア開発には、平田裕也氏、樋口雄紀氏、杉山詩歩氏、杉山月渚氏、生駒創氏にご協力いただいた。





図 4.4: ジャがいものデータ取得の様子 (広島県東広島市にて)。

#### 4.3.3 データセットの構成

本研究で構築したデータセットは、レモン、みかん、じゃがいもの3種の作物を対象として収集した。撮影方向数については、レモンが本来5方向を予定していたものの、機械搬送時のトラブルにより1カメラが故障し、最終的に4方向となった。みかんとじゃがいもについては、それぞれ5方向から撮影している。サイズや等級のクラス分け基準は、農協が定めた規格に準拠し、表面の傷や変色、形状の乱れといった外観上の要素を熟練した農家が判定している。

また、撮影した作物数は、レモン7363個、みかん4323個、じゃがいもで2285個に上り、クラスごとの内訳は表4.1、表4.2、表4.3に示すとおりである。撮影条件に関しては、露出や照明の強度を一定に保ち、外光の影響を極力遮断するように管理を行った。なお、みかんのデータセットに一部欠損が見られるのは、撮影データの保存時に破損ファイルが生じたためであり、今後は機材の予備確保やバックアップ体制の強化などを検討している。



構築したデータセットの主な特徴を以下に示す。

- 対象作物と撮影方向数:

- － レモン：4 方向 (本来は 5 方向予定だが機械不調のため 1 カメラ分を除外)
- － みかん：5 方向
- － ジャがいも：5 方向

- クラス分け基準:

- － レモン
  - \* サイズ：S, M, L, 2L, 3L
  - \* 等級：秀, 優, 良
- － みかん
  - \* サイズ：2S, S, M, L, 2L, 3L, 4L
  - \* 等級：1 級品, 2 級品, 規格外
- － ジャがいも
  - \* サイズ：M, L, 2L
  - \* 等級：秀, 優, 規格外

サイズの基準としては、農協の既存規格に準拠している。等級評価はその土地の選果基準をもとに農家に行っていただいた。図 4.5 に、実際に取得した作物のデータ例を示す。レモン、みかん、ジャがいもそれぞれ、秀品 (みかんは 1 級品) と規格外品 (レモンは良品) のうちの一つである。目視でも、規格外品にはレモン・みかんでは傷が、ジャがいもには割れがあることがみて取れる。

- 撮影枚数・分布:

- － レモン：7363 個
- － みかん：4323 個
- － ジャがいも：2285 個

クラスごとの撮影枚数は表 4.1 から 4.3 に示すとおりである。撮影条件に関しては、露出や照明の強度を一定に保ち、可能な限り背景や環境光の影響を排除するよう管理した。

なお、レモンの撮影方向が 4 方向に留まった理由は、撮影機械搬送時のトラブルにより 1 方向のカメラが故障したためである。また、みかんのデータセットのクラスにいくつか欠陥が見られるのは、データが破損したためである。今後は機材の予備の用意・綿密なバックアップなどの対策を検討している。

#### 4.3.4 データセットの価値と意義

本研究で構築したデータセットは、学術的な研究から産業応用に至るまで幅広い価値を持つと考えられる。まず、従来の農作物画像が一方のみの撮影に偏りがちだったのに対して、本データセットではレモン、みかん、ジャがいもについて 4～5 方向からのマルチビュー撮影を実現してお



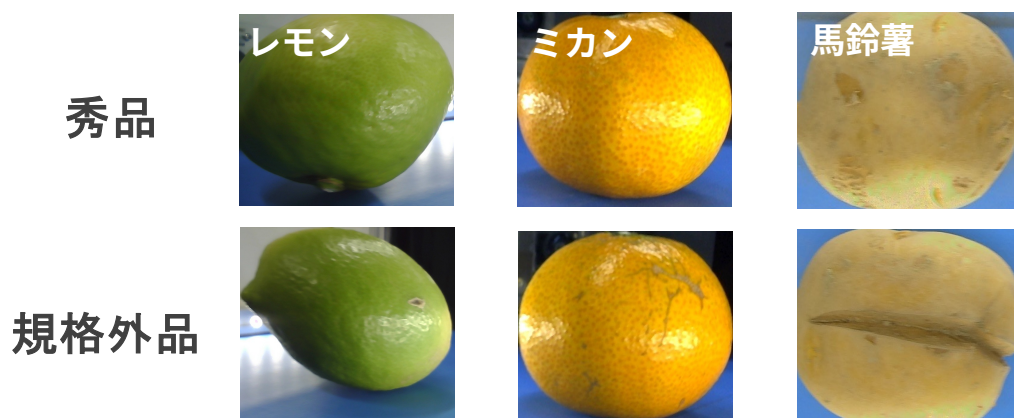


図 4.5: 取得したレモン，みかん，じゃがいものデータ例．最も等級が高いものと低いものをそれぞれ一つずつ示す．

表 4.1: レモンのサイズ・等級分布

サイズ	等級			
	秀	優	良	合計
S	109	97	109	315
M	1415	1285	798	3498
L	701	696	649	2046
2L	217	340	411	968
3L	43	120	373	536
合計	2485	2538	2340	7363

り，単一視点では把握しづらい形状の特徴や微妙な外観差異をより多角的に記録している．さらに，複数の品目を同一環境下で撮影しているため，汎用的な等級選別モデルの開発や異なる作物間でのドメイン適応研究にも活用しやすく，モデル性能の客観的評価や比較検討を行う上でも有用である．

また，本データセットは，既存の公開データセットと比較しても，多視点撮影と等級評価ラベル（サイズと等級の組み合わせ）が充実しているという点で差別化が図られている．サイズ選別と農家による等級判断の両方を反映したラベル情報が整理されていることで，研究コミュニティにとって新たなベンチマークとして活用できる．さらに，ディープラーニングを用いた高精度のモデル構築に必要なサンプル数が確保されており，国内農業の労働力不足を解消するための省人化・自動化技術の研究開発を後押しする資源としても期待される．

以上のように，多方向・多品種・大規模かつ生産現場の実態に即した本データセットは，農作物の品質分析や自動選別技術に関する研究開発で大きな役割を果たすと考えられる．現在はサイズ



表 4.2: みかんのサイズ・等級分布

サイズ	等級			
	1 級品	2 級品	規格外	合計
2S	-	255	335	590
S	926	641	149	1716
M	330	107	54	491
L	-	-	447	447
2L	160	-	-	160
3L	617	33	71	721
4L	149	2	47	198
合計	2182	1038	1103	<b>4323</b>

表 4.3: ジャがいものサイズ・等級分布

サイズ	等級			
	秀	優	規格外	合計
M	323	217	413	953
L	242	232	419	893
2L	167	90	182	439
合計	732	539	1014	<b>2285</b>

と等級のラベリングが中心となっているが、将来的には、作物表面の傷や欠陥部位の詳細ラベリングを行うことで、さらに高度な品質評価や幅広い用途への応用を実現できる可能性があると考えられる。

## 4.4 実験

本研究では農作物の選別工程の自動化のため、構築したデータセットを訓練に用いることで、等級選別の精度の検証を行う。データセットはサイズ、等級と二つの属性を持っているが、本実験ではサイズは関係なしに、等級ごとの分類精度の検証を行う。レモン、ジャがいもに関しては全てのサイズを用いた。みかんに関してはデータに欠損があるため、S と M サイズのみを用いた。多視点から得られた各作物の画像を入力とし、等級分類を行うために ResNet-18 モデルを採用した。ResNet は He らによって提案された畳み込みニューラルネットワークであり [6]、最も一般的に用いられる深層分類モデルである。実装および学習には PyTorch フレームワークを用い、下記のハイパーパラメータを設定した。

- 学習率: 0.0005 (学習の進行に伴いスケジューラで減少)



表 4.4: 作物ごとの分類性能 (多視点データを使用).

作物	Accuracy (↑)	MCC (↑)
レモン	0.853	0.780
みかん	0.897	0.809
じゃがいも	0.843	0.756

- 最適化アルゴリズム: Adam
- 損失関数: クロスエントロピー損失
- エポック数: 30
- バッチサイズ: 32

多視点撮影によって得られた複数画像を全てモデルの入力として用いることで、単一方向からの撮影では捉えきれない作物の形状や表面の特徴を捉えることができ、より高い精度での分類が可能となる。本モデルでは、多視点撮影によって得られた複数画像 (レモンは 4 視点、みかんとじゃがいもは 5 視点) の平均画像を ResNet-18 に入力する方式を採用している。複数画像を個別にモデルに入力し、出力された各視点の特徴量またはクラススコアを平均プーリングによって統合する方式も考えられたが、複数回の推論が必要となり計算コストが大きいため今回は採用しなかった。

構築したデータセットについては、作物ごとに訓練データ 70%、検証データ 15%、テストデータ 15% の割合で分割した。各作物 (レモン、みかん、じゃがいも) についてそれぞれ独立にモデルを学習し、テストデータを用いて性能評価を行っている。分割にあたっては、サイズや等級の偏りが極端にならないようランダムに抽出し、学習時にはデータ拡張として左右上下反転をランダムに施した。

各作物ごとに学習したモデルの性能を評価するため、Accuracy と MCC (Matthews Correlation Coefficient) を用いた。Accuracy はクラス全体の正解率を示す指標であり、分類タスクの最も基本的な評価方法である。一方、MCC はクラス間のサンプル数に偏りがある場合でも頑健に評価が可能な指標とされ、-1 から 1 までの範囲で相関を評価する [26]。今回のデータセットはサイズや等級のクラス分布が一樣ではないため、MCC を併用することでより総合的なモデルの性能分析が行えると考えた。

## 4.5 実験結果と考察

今回の実験では単視点データとの比較は行わず、各作物間での性能差に着目して考察を進める。表 4.4 に、レモン、みかん、じゃがいもごとに多視点データを用いて学習したモデルの Accuracy および MCC を示す。**みかん**での Accuracy および MCC が最も高く、**レモン**は中間程度、**じゃがいも**はやや低い値となっている。**みかん**は 0.897 という高い Accuracy と 0.809 という高 MCC



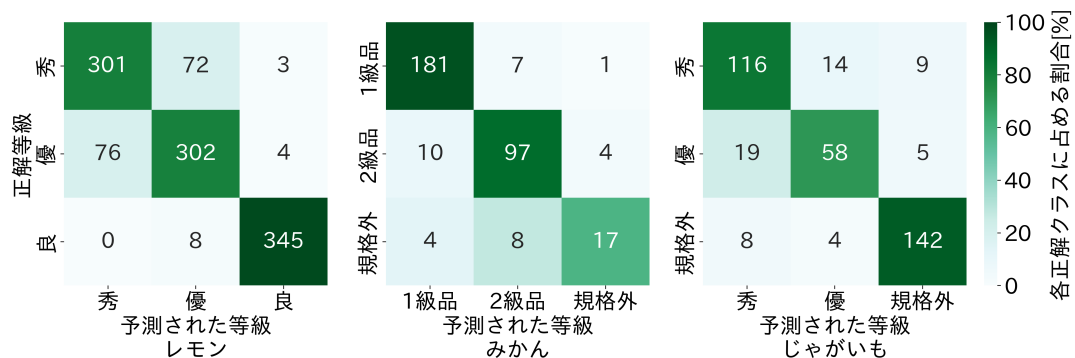


図 4.6: 各作物のテストデータに対する混同行列.

を達成しており、これは多視点情報によって表皮の微妙な色合いやサイズ感の違いがより適切に捉えられているためと考えられる。一方で、**レモン**は5視点ではなく4視点のみの撮影となった影響もあり、Accuracyは0.853にとどまったが、MCCが0.780と比較的高い。これは、**みかん**は規格外品のデータが1級品、2級品と比べて少ないのに対して、**レモン**は等級のデータ数のバランスが取れているためである。**じゃがいも**は形状やサイズ差が明瞭な半面、皮の色や微細な傷を識別する際に誤差が生じやすく、Accuracyが0.843、MCCが0.756という結果になったと考えられる。

さらに、図 4.6 に各作物のテストデータに対する混同行列を示す。**レモン**の場合 (左図)、クラス間の誤分類は主に**秀品**と**優品**との間で多く見られ、それぞれ72、76の誤分類が発生している。これは、実際には等級の異なる個体を複数の視点から撮影しても、モデルが決定境界を十分に学習できなかった可能性を示唆している。**みかん**の場合 (中央図)、対角成分 (181, 97, 17) が大きく、誤分類数が比較的小さいことが高精度につながっている。特に**1級品**では誤分類が7件と1件と少なく、明確な外観特徴をモデルが安定して把握できていると考えられる。**じゃがいも** (右図) については**秀品**と**規格外品**の正解数が高いものの、**優品**に対しては誤分類が相対的に目立ち、等級が近い個体同士の判定が曖昧になっていると推測される。

以上の結果から、多視点データを活用したモデルは、単一視点では捉えきれない形状や表面の特徴を統合することで、全体的に高い分類性能を実現できることが確認された。一方で、各作物特有の外観要素 (表皮の微妙な色合い、サイズが近い個体の境界、傷やへこみなど) の違いによって、モデルの学習難易度や誤分類のパターンにも差が生じている。

## 4.6 未来の研究方向性

本研究では、多視点画像を活用した農作物の等級分類モデルを構築し、レモン・みかん・じゃがいもを対象に有効性を示した。しかし、実運用に向けてはいくつか解決すべき課題や追加の研究展開が考えられる。まず、作物を検出する段階で精度の低下が見られる場合があり、その影響で一部の作物画像が正常に切り出せず、データセットに欠損や誤検出が含まれている点が懸念される。



たとえば背景とのコントラストが十分でない状況や、撮影時に作物がぶれたり重なったりする場合には、データの欠損が生じやすい。今後は、本研究で得られたデータを精査し、撮影環境のさらなる安定化や、物体検出アルゴリズム (例: YOLO や Mask R-CNN) の改良によって、誤検出と欠損データの発生を抑制し、より正確な作物領域抽出ができるようにする必要がある。

また、本研究のようにディープラーニングを用いたモデルは、推論過程がブラックボックス化しやすいという課題が指摘されている。農家や選果場での実運用を想定した場合、収益性の観点から「なぜこの作物が該当クラスと判断されたのか」を説明できる手法が求められる。具体的には、Grad-CAM のような可視化技術や注意機構の導入による領域ハイライト、モデルの学習プロセスを解釈しやすくする説明可能な AI のフレームワークの検討などが挙げられる。これにより、現場担当者がモデルの判断に納得感を持ちやすくなり、必要に応じてラベル修正やデータ増強の方針決定を行いやすくなると期待される。

さらに、表面の傷や欠陥の箇所を検出する技術についても検討が必要である。本研究では深層学習モデルの特徴量抽出によって等級分類を行ったが、運用実態の中では表皮の傷や変色の程度といった、古典的な画像処理で検出できる要素が品質評価に大きく影響することが多い。したがって、今後は画像解析技術を高度化し、傷の位置や形状、色の画素値も入力として含めたモデル設計が考えられる。また、そのような検出結果を用いて、等級付けの根拠を可視化したり、過去の類似ケースと比較してシステムが自動的に判断基準を更新したりする仕組みが考えられる。

最後に、異常検知技術の応用により、一般的なクラス分類では対応しきれない欠陥や特殊な病害などを発見する可能性がある。例えば、一定の特徴分布から外れたサンプルをスコアリングで洗い出すことで、これまで経験のないような変色や内部腐敗などを検出し、品質管理に役立てることが期待される。特に、目視検査することが難しい病気や傷も存在するため、こうした異常検知技術は省人化の大きな武器になると考えられる。

## 4.7 結論

本研究では、深刻化する国内農業の労働力不足を背景に、レモン・みかん・じゃがいもを対象とする多視点データセットを新たに構築し、ディープラーニングを用いた等級分類モデルの開発と評価を行った。従来の単一視点撮影が主体のデータセットとは異なり、本研究で構築したデータセットは 4~5 方向からの多視点撮影を行い、サイズや外観品質 (等級) といったラベルを整備することで、形状や傷、色むらなどの多様な外観情報をより正確に反映している。これにより、単一視点では把握しづらい作物の微妙な差異や、撮影条件のばらつきによる誤判定のリスクを軽減し、汎用性の高い研究基盤を提供できる。実験では、構築した多視点データを用いて ResNet-18 モデルを学習させることで、高い精度を示すことが明らかになった。各作物固有の課題も存在したものの、総じて多視点活用は外観選別の自動化において有望なアプローチであることが示唆される。しかしながら、作物領域の検出精度が十分でない場合の誤差や、ディープラーニングの推論過程がブラックボックス化しがちな点など、実運用に移行するうえで考慮すべき課題が顕在化したのも事実である。今後は、物体検出アルゴリズムの高精度化や、説明可能な AI (XAI) の導入、傷や欠



陥を詳細にラベリングしたデータ拡張などを積極的に進める必要がある。また、複数品目を同時に扱う場面ではドメイン適応や転移学習の活用が重要となり、多視点画像から抽出した形状の情報をどのように統合しモデル入力するかが、さらなる高精度化と汎用化に重要であると考えられる。



## 第 5 章

# 自動 AI 選果システム実現に向けた現在の課題と未来の研究方向性

本研究では、画像処理や深層学習技術を活用した自動選別の有用性を示してきたが、こうした技術が実際の農業現場に普及するためには、実地導入を仮定した際の課題への対処と要件整備が欠かせない。特に、国内の農家や農協の関係者へ直接インタビューを行った結果、設備投資コストの高さや選別基準の多様化、さらには多品目への対応といった点が導入を阻む大きな障壁として浮かび上がった。本章では、それらの課題を具体的に整理するとともに、本研究成果の延長上でどのようなアプローチを取るべきか、未来の研究方向性を論じる。

### 5.1 インタビューを通じて判明した課題

国内で初めて AI を導入した「JA 三ヶ日みかん選果場 (柑橘販売課, 柑橘指導課)」およびその AI 機器の開発を委託された「シブヤ精機株式会社 (浜松本社)」, さらには柑橘農家 2 軒へのインタビューを通じて、以下の三つの主要な障壁がユーザーにとっての課題であることが判明した。

国内で初めて AI を導入した JA 三ヶ日みかん選果場 (柑橘販売課, 柑橘指導課) や、その AI 機器を開発したシブヤ精機株式会社, そして柑橘農家へのヒアリングによって、ユーザー視点から見た三つの主要な課題が明確となった。第一の課題は「設備投資の困難さ」である。大規模な選果工場を建設する事例は一部に存在するが、一般的な地域や個人農家では投資可能な金額が限られ、高額なシステム導入は事実上困難となっている。補助金を含めても 500 万～600 万円以下の設備しか選択肢に入らない場合が多く、毎年数億円単位で選別工程に投資する大規模農協でも、組織文化や IT リテラシーの観点で新技術導入が進まずにいるという指摘もなされた。

第二の課題は「選別基準の固定化問題」である。国内の農家が扱う作物や品種は多岐にわたり、収穫時期や販路によっても外観の要求水準が変化するため、品目ごとに複数のランクを組み合わせた複雑な基準が用いられることが珍しくない。しかし、既存の選別機や AI システムは多様な基準を想定していないケースが多く、一律化された設定では地域や農家の実態に合わないことがある。JA 三ヶ日の例では、AI 品質選別施設を利用する代わりに、2～3 割の農家が自前の手選別を



続けていることが象徴的である。農協の統一基準によって販路や価格決定権が制限されるリスクを敬遠し、それぞれの農家が求める柔軟な基準を活かしたいというニーズがあるのだ。

第三の課題は「多様な作物品種ごとの導入が必要」な点である。単一の品種に特化して大規模投資を行う地域もある一方で、多くの地域では少量多品目の栽培が行われ、個々の作物に膨大なデータを用意して学習させる余力はない。品種が増えるほどデータセットの維持やアップデートの負担が高まり、システム運用そのものが複雑になる。こうした点から、AI 選別機の導入を検討する農家や生産者組織があっても、最終的には従来の手選別に頼ってしまうという現状が浮かび上がった。

## 5.2 実地導入のために追加で必要な要件

これらの課題を踏まえると、実地での導入・運用を成功させるには、単に高精度な AI モデルを開発すればよいというものではなく、コストやカスタマイズ性、品目拡張性を考慮した総合的なシステム設計が求められる。第一に「低コスト化」が不可欠となり、500 万円以内で導入可能な小型選別機の開発が技術普及の要となる。この価格帯を実現しつつ、高価な大型装置に引けを取らない程度の精度と処理速度を確保するためには、撮影や搬送装置の簡素化や、計算資源を抑えたアルゴリズム設計が必要である。

第二に「選別基準の柔軟なカスタマイズ」が重要である。農家ごと、あるいは地域ごとに異なる外観基準や傷の許容度合いを、ユーザーフレンドリーな UI で簡単に設定・変更できる仕組みがなければ、導入後に運用面の不便が生じ、ユーザーが AI システムを忌避する要因になりうる。たとえば、図 5.1 のようなグラフィカルインタフェースを用い、明度や傷サイズ、変形率などをパラメータとして指定・調整する機能を搭載すれば、農家自身が選別要件をリアルタイムで変更可能となり、個々の出荷先や時期に合わせた柔軟な選別が行える。

第三に「多品目対応」の側面では、汎用的な基盤モデルの構築が課題となる。異なる品目や品種でも共通する外観特徴（たとえば、色空間の扱いや形状表現など）を学習しておき、新しい作物が追加された場合は転移学習を活用して少量のデータで素早く適応できる仕組みが有効である。本研究で整備したような多方向からの大規模データセットは、こうした基盤モデルの汎化性能を高めるための大きな資源となる。さらに、ハードウェア面でも品目ごとのサイズや形状差に対応できるよう、アクチュエータや搬送ラインの調整を行いやすいモジュール化が望まれる。たとえば、果実のサイズに合わせてガイドや排出口を交換できる設計にするなど、物理面とソフトウェア面を一体的に考慮することが必要である。

## 5.3 今後の展望と研究方向性

以上の課題と要件を総合すると、今後は「低コスト・モジュール型」の AI 選別機を開発し、農家自身が必要とする選別基準を手軽に再学習・再設定できる環境を整えることが鍵となる。具体的には、第一に、今回のように多視点で撮影された大規模データセットをさらに拡充し、点群情報





図 5.1: 選別機のためのユーザーフレンドリーな UI デザイン。

なども組み合わせたマルチモーダル学習を進めることで、少量データでも高精度に適応できる基盤モデルを完成させる。第二に、農家向けの操作画面をより直感的に設計し、タッチパネルやスマホアプリから選別要件を修正できる仕組みを整える。第三に、機械のメンテナンス性や拡張性を高めるために、コンパクトなカメラ・照明・アクチュエータユニットを複数組み合わせるモジュール方式を採用し、品目や出荷量の変化に柔軟に対応可能とする。

これらの取り組みを進めることにより、農家や小規模生産組織でも手が届きやすい価格帯の自動選別機が普及し、選別基準の差異や多品目生産に伴う負担を軽減できると期待される。特に、少子高齢化に直面する地方の農業地域では、現場作業の大幅な省人化と効率化が急務となっている。本研究で示したように、ディープラーニングを駆使した外観評価モデルは一定の効果を発揮できるが、そこから一歩進んで、「どのように導入し、どのように運用すれば現場で受け入れられるか」を追求する段階にきていると言える。実際に導入を決める農家や農協にとっては、機器の価格だけでなく、カスタマイズ性やアフターサポート、転移学習の容易さなど複合的な要因を吟味する必要がある、研究開発のさらなる発展が望まれる。

総じて、AI 選果システムの実現には技術的要素と社会的要素の両面からのアプローチが求められる。高精度なモデルと多品目対応のデータセットを整備するだけでなく、低コストかつモジュール型のハードウェア設計やユーザーフレンドリーなインターフェース開発、地域や農家の事情に合わせた柔軟な運用モデルが不可欠となる。こうした要件を踏まえつつ、今後は撮影環境の最適化や 3D 情報の活用、異常検知の高度化など多角的な研究を継続していくことで、自動 AI 選果システムの普及と日本農業全体の活性化に貢献できると期待される。



## 第 6 章

# 結論

本論文では、国内農業の深刻な労働力不足を背景に、作物選別工程の自動化を目指して二つの研究アプローチを示し、さらにはインタビュー調査から得られた現場の課題や要件分析を通じて、実地導入に向けた展望と課題を総合的に論じた。まず、第 3 章では三次元点群データを活用し、みかんを対象に画像と点群の両方を入力とするマルチモーダルモデルを構築することで、二次元画像単体では捉えきれない外観上の微細な形状差異を補完できることを示した。結果として、画像＋点群を同時学習させるアプローチが極めて高い分類精度を達成し、三次元情報の活用が品質選別の高精度化をもたらす可能性を示唆した。一方で、三次元スキャンに関するハードウェアコストの増大やノイズ除去、点群単体のモデルが伸び悩むなど、導入・運用面の課題も明確になった。

次に、第 4 章では多視点撮影が可能な独自の撮影機械を開発し、レモン・みかん・じゃがいもを対象とする大規模データセットを構築した。従来は単一視点の画像に偏りがちだった既存の公開データセットに対し、本研究で構築したデータセットは 4～5 方向からの画像を大量に収集し、サイズや等級のラベルを付与することで、より実用的かつ汎用性の高い学習環境を実現した。実験の結果、複数方向からの情報を活用することで、単一視点では捉えきれない傷や色むら、形状の歪みなどをより正確に把握でき、3 種の作物に対していずれも高い分類性能を得られることを確認した。これは、多視点によるデータ拡充が生産現場の複雑な品質基準を柔軟に把握するために重要であることを示す。

さらに、AI 選果システムの導入を検討している農家や農協、システム開発企業へのインタビューを通じて、設備投資の負担や選別基準の多様化、多品目への展開の難しさといった、技術面以外のボトルネックが明らかとなった。これらは、高精度なモデルを開発するだけでは解決できず、低コスト化・操作性・柔軟な基準設定といった要素を含む総合的なソリューションが求められることを示している。本研究で収集した大規模・多視点データセットや画像＋点群の併用モデルは、そうした課題に対して一つの解決策を提示できるが、実地での導入を進めるためには、安価なハードウェア設計やユーザーフレンドリーなインターフェースの導入、転移学習による多品目対応などをさらに推し進める必要がある。

本研究の成果は二次元画像と三次元点群、あるいは多視点撮影によるデータ拡張が、農作物の外観品質選別において有効な手段であることを示し、国内農業における AI 活用の可能性を拓くも



のである。同時に、現場との連携を強化し、コスト面・操作面・運用面を総合的に改善していく道筋を明らかにしたことにも重要な意義がある。今後は、より安価な 3D センシング技術の普及や、センサネットワーク化された選別ライン、さらには異常検知や傷部位の可視化といった高度な応用を組み合わせることで、AI 選果システムが一層精密かつ柔軟に進化していくと期待される。これらの取り組みが広く普及すれば、日本の農業が抱える人手不足を補いつつ、生産効率を高め、世界市場への競争力強化にもつながると考えられる。本研究はその第一歩として、学術的にも産業的にも有用な基盤データセットとモデルの検証事例を提示した。今後の研究では、本論文で提起した課題解消に向け、現場との連携を図りながらより高度な AI 選別技術の確立と実用化を進めていく。



## 参考文献

- [1] 農林水産省. 2020 年 農林業センサス, 2020.
- [2] 農林水産省. 令和 5 年農業経営をめぐる情勢について, 2023.
- [3] 農林水産省. 令和 4 年農業経営体の経営収支, 2023.
- [4] An introduction to convolutional neural networks. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, Vol. 10, pp. 943–947, 2015.
- [5] Piotr Rybacki, Janetta Niemann, Samir Derouiche, Sara Chetehouna, Islam Boulaares, Nili Mohammed Seghir, Jean Diatta, and Andrzej Osuch. Convolutional neural network (cnn) model for the classification of varieties of date palm fruits (phoenix dactylifera l.). *Sensors*, Vol. 24, No. 2, p. 558, 2024.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CVPR*, 2016.
- [7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.
- [8] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *CVPR*, 2016.
- [9] Yuhang Fu, Minh Nguyen, and Wei Q. Yan. Grading methods for fruit freshness based on deep learning. *SN Computer Science*, Vol. 3, p. 264, 2022.
- [10] Bin Zheng and Tao Huang. Mango grading system based on optimized convolutional neural network. *Mathematical Problems in Engineering*, Vol. 2021, pp. 1–11, 2021.
- [11] Juan Daniel Arango, Benjamin Staar, Adil Maqsood Baig, and Michael Freitag. Quality control of apples by means of convolutional neural networks-comparison of bruise detection by color images and near-infrared images. *Procedia CIRP*, Vol. 99, pp. 290–294, 2021.
- [12] Ang Wu, Juanhua Zhu, and Taiyong Ren. Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network. *Computers & Electrical Engineering*, Vol. 81, p. 106454, 2020.
- [13] Yu Yang, Zhenfang Liu, Min Huang, Qibing Zhu, and Xin Zhao. Automatic detection of multi-type defects on potatoes using multispectral imaging combined with a deep learning model. *Journal of Food Engineering*, Vol. 336, p. 111213, 2023.



- [14] Wenqi Zhou, Chao Song, Kai Song, Nuan Wen, Xiaobo Sun, and Pengxiang Gao. Surface defect detection system for carrot combine harvest based on multi-stage knowledge distillation. *Foods*, Vol. 12, No. 4, p. 793, 2023.
- [15] José Mendoza-Bernal, Aurora González-Vidal, and Antonio F Skarmeta. A convolutional neural network approach for image-based anomaly detection in smart agriculture. *Expert Systems with Applications*, Vol. 247, p. 123210, 2024.
- [16] Yisen Liu, Songbin Zhou, Zhiyong Wan, Zefan Qiu, Lulu Zhao, Kunkun Pang, Chang Li, and Zexuan Yin. A self-supervised anomaly detector of fruits based on hyperspectral imaging. *Foods*, Vol. 12, No. 14, p. 2669, 2023.
- [17] Yulan Guo, Hanyun Wang, Qingyong Hu, Hao Liu, Li Liu, and Mohammed Bennamoun. Deep learning for 3d point clouds: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 43, pp. 4338–4364, 2019.
- [18] Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. *CVPR*, 2017.
- [19] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *NeurIPS*, 2017.
- [20] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *NeurIPS*, 2012.
- [21] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip Torr, and Vladlen Koltun. Point transformer. *ICCV*, 2021.
- [22] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *NeurIPS*, 2017.
- [23] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. *ICLR*, 2021.
- [24] Weihao Weng and Xin Zhu. U-net: Convolutional networks for biomedical image segmentation. *MICCAI*, 2015.
- [25] Xu Ma, Can Qin, Haoxuan You, Haoxi Ran, and Yun Fu. Rethinking network design and local geometry in point cloud: A simple residual mlp framework. *ICLR*, 2022.
- [26] Davide Chicco and Giuseppe Jurman. The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC genomics*, Vol. 21, No. 1, pp. 1–13, 2020.



## 発表文献

- [1] 水谷航悠, 山崎俊彦. 3次元点群情報を考慮した農作物の外観品質選別. 画像符号化シンポジウム/映像メディア処理シンポジウム (PCSJ/IMPS), 2024 (ポスター発表).
- [2] 水谷航悠, 山崎俊彦. 農作物選別作業の自動化に向けたデータセット構築と品質分類モデルの開発. 画像工学会 (IE), Feb. 2025 (発表予定).



# 謝辞

本研究を進めるにあたり、多くの方々のご指導とご協力を賜りましたので、ここに深く感謝の意を表します。まず、山崎研究室で過ごした修士課程の2年間は、瞬く間に過ぎたと感じるほど充実した日々でした。山崎俊彦教授には、毎週ディスカッションの場を設けていただき、研究内容のみならず、プレゼンテーションや英語学習に至るまで幅広いご指導とサポートをいただきました。おかげさまで、今年1年間を有意義なものにできたと心より感謝しております。来年度からも引き続き、ご指導・ご鞭撻のほどよろしくお願い申し上げます。また、相澤清晴教授、山肩洋子教授、松井勇佑講師には、ミーティングや発表のリハーサルなどの折に貴重なコメントを数多く頂戴しました。これらのアドバイスは研究を深めるうえで大変参考になりました。併せて御礼申し上げます。特任専門職員の江川綾様には、研究や大学生活に関わるさまざまな事務手続きでお世話になりました。迅速かつ丁寧にご対応いただき、心強かったです。ありがとうございました。研究室の先輩方には、常に温かく接していただいたうえ、発表や研究手法など多岐にわたるアドバイスをいただきました。これらが自分の成長を後押ししてくださったと感じております。また、同じ志を持つ優秀な同期、山西博雅君や Ye Xiaotong 君の存在は、研究を続ける大きな支えとなりました。皆様に心より感謝いたします。また、本研究の実施において、アイデア出しやコーディング補助、論文の構成案作成および文章推敲のサポートに ChatGPT を活用しました。

本研究においては、データ取得の協力やインタビューへのご対応など、多くの企業・団体の方々にご尽力を賜りました。岡宗農園の岡宗俊介氏には「はるか」の画像と点群データの取得のため、作物を提供してくださいました。広島県庁の皆様、とりわけ長戸玄氏、東山聖恵氏、広島県果実農業協同組合連合会の皆様には、貴重なお時間を頂戴し、レモン・みかんの画像データ収集のご協力をいただきました。中市後久夫氏には、じゃがいもの画像データ収集のご協力をいただきました。さらに、シブヤ精機株式会社様、ミヤシタ果樹園様、小田農園様、サンエイ株式会社様、株式会社ニレコ様、ダイヤモンド十勝株式会社様、十勝とやま農場様、JA かながわ西湘様、JA いちかわ様、JA 豊橋様、JA ありだ様、そして豊橋市様にも、インタビューや各種調整に多大なるご協力をいただきました。皆様のお力添えがなければ、本研究をここまで進めることは叶わなかったと痛感しております。この場を借りて、改めて深く御礼申し上げます。

また、本研究の遂行にあたり、NEDO の「研究開発型スタートアップの起業・経営人材確保等支援事業／ディープテック分野での人材発掘・起業家育成事業 (NEP) ／躍進コース」に採択いただくとともに、IPA の「IPA 未踏アドバンスト事業」にも採択いただきました。これらのご支援により研究資金を確保することができ、大いに助けられました。心より御礼申し上げます。



最後に，大学での学びを支え，ここまで通わせてくれた父と母に心より感謝いたします．

2025 年 1 月 23 日