

## CNN を用いた農作物の収量予測と Grad-RAM による可視化の検討 Crop Yields Prediction with CNN and Visual Explanation with Grad-RAM

下村 真生<sup>†</sup> 中畝 誠<sup>‡</sup> 山田 融<sup>‡</sup> 塚澤 和憲<sup>‡</sup> 中村 和幸<sup>†</sup>  
Masaki Shimomura Makoto Nakaune Toru Yamada Kazunori Tsukazawa Kazuyuki Nakamura

### 1 はじめに

農作物の収量予測は、生産量安定化および収益安定化の観点から重要であるが、現在は生産者の経験による推測で判断をしている。しかし、正確な予測には植物の生育状況や生育環境のデータ解析が必要であり、近年は人工知能技術を用いた容易かつ定量的な予測の需要が高まっている。現在、オリーブについては樹木の画像から果実を抽出し、その数から収量を予測した報告 [1] やトマトについて、その果実を撮影し、Regional Convolutional Neural Network (RCNN) や Single Shot Detector (SSD) などを用いて収穫量を予測した報告 [2] などがあるが、きゅうりを対象とした研究は十分でない。

本研究では、画像と生育・環境データを入力とする CNN を用い、1 週間後の収量予測を行う。その際、実地での応用時に生産者が簡単に使用できることを目的として、用いる画像は区画全体を捉えた可視光画像とした。さらに、Grad-CAM を回帰モデルに適用できる形に変更した Grad-RAM による予測根拠の視覚的説明を行う。

### 2 実験手法

#### 2.1 学習条件

画像データは、2018 年 9 月から 2018 年 12 月および 2019 年 4 月から 2019 年 7 月の 2 期間にきゅうりを栽培した施設 2 棟に対し、各区画を撮影した画像 894 枚 (データ不足を補うため、各区画 10 枚撮影) を用いる。また数値データとして、画像撮影日における各ハウス内の気温 [°C]、日射量 [kW/m<sup>2</sup>] と各区画代表株の雌花数、開花節数、開花節長、開花節径の観測値 447 時点分を用いる。なお、気温と日射量は撮影日から遡った過去 7 日間の毎時 00 分の合計値を使用した (以下それぞれ 7 日間累積気温および 7 日間累積日射量)。実験に際し、予測日として 2018 年 10 月 22 日と 2019 年 6 月 22 日の 2 日、計 30 区画分を設定し、この 2 日を除いたデータのうち 10% (42 区画分) を評価用データ、残る 375 区画分を学習データとした。

本実験で使用した、画像データと数値データの両者を用いた CNN モデル (model C) を図 1 に表示した。画像データを入力とし ResNet50 をベースとした CNN (model A) と、数値データを入力とした NN (model B) を結合し、収量を出力としている。なお、model B は 6 種の数値データを入力とし、中間層はなく、出力サイズを 1 とした。また、いずれのモデルも出力層の活性化関数は恒等関数とした。損失関数は平均二乗誤差、最適化手法は Adam ( $\beta_1 = 0.9, \beta = 0.99$ ) を用いる。ミニバッチのサイズは 32 とし、画像は 512 × 512 にリサイズし、画素値を 1/255 とした。また、各エポックでミニバッチを生成する際にデータ拡張として、ランダムに水平方向への反転

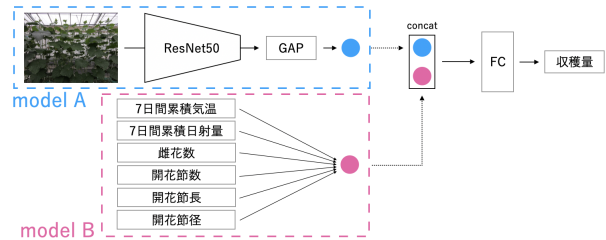


図 1: CNN モデル (この全体が model C)

した。この条件で学習した model A と model B の重みを初期値として、model C を同じ条件下で学習した。

#### 2.2 Grad-RAM : 予測根拠の視覚的説明

収量予測の根拠を視覚的に説明する手法として Grad-RAM を用いた。画像分類に関しては Class Activation Mapping (CAM)[3] およびその定義の一部を一般化することで、より多くの CNN モデルへ適用可能な形へと変更した手法である Grad-CAM[4] の適用事例が多い。また、回帰モデルに関しても Regression Activation Map (RAM)[5] 等の適用事例がある。RAM は CNN モデルにおける最後の畳み込み層に対して Global Average Pooling を施し、直後に出力層へ繋がることを要求している。しかし本研究で使用した model C のように、出力層付近がこの構造でない CNN モデルに対して、視覚的説明を必要とする場合も多い。そこで Grad-CAM 同様に多くの CNN モデルへ適用可能な手法として本研究では Grad-RAM を提案する。Grad-CAM の定義についてクラス数を 1、活性化関数を恒等関数としたものが Grad-RAM である。

出力層の出力  $Y$  と特徴量マップ (最後の畳み込み層) の  $k$  番目のチャンネルにおける位置  $(x, y)$  の活性  $A_{xy}^k$ 、特徴量マップの面積  $z$  から得られる重み:

$$w_k = \frac{1}{z} \sum_x \sum_y \frac{\partial Y}{\partial A_{xy}^k}, \quad (1)$$

と顕著量マップ  $L_{xy}$ :

$$L_{xy} = \text{relu} \left( \sum_k w_k A_{xy}^k \right), \quad (2)$$

を算出する。ここで  $\text{relu}(\cdot)$  は ReLU 関数で  $\text{relu}(x) = \max(0, x)$  を表す。また、 $L_{xy}$  を入力画像のサイズまでアップサンプリングしたものを  $\tilde{L}_{xy}$  とする。なお、アップサンプリング手法としては OpenCV[6] で実装された Lanczos 法を利用した。この  $\tilde{L}_{xy}$  を入力画像に重ね合わせることで、ヒートマップを用いた視覚的説明を得られる。

<sup>†</sup> 明治大学 Meiji University

<sup>‡</sup> 埼玉県農業技術研究センター Saitama Agricultural Technology Research Center

表1: 3モデルのテスト結果

モデル	A(画像)	B(数値)	C(画像+数値)
MAE	1055.5	654.2	583.5

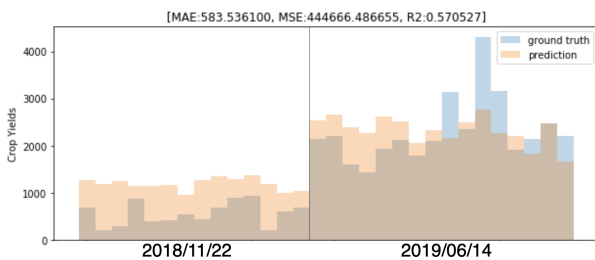


図2: model C の予測結果

### 3 結果

#### 3.1 CNNモデルの学習

model A, B, Cによる予測テスト結果は表1である。評価には平均絶対誤差を利用した。画像のみを用いたmodel AではMAEが1000を超えたが、数値データを合わせることでmodel CではMAEが583.5となり、MAEが45%程度低下した。図2ではmodel Cの区画ごとの予測結果を示した。予測日の真の収穫量を青、model Cによる予測結果を橙で表示している。なお、学習時model Bは1区画1データに対し、model A, Cは1区画10データで学習したが、テスト時はmodel A, Cで1区画10データの予測結果の平均をとることで、いずれのモデルも1区画1予測値としてテスト精度を算出した。冬季はおよそ1000g程度であり、夏季はおよそ2500g程度であることから、季節間の収穫量の差は見取れるが、同じ日の区画間で十分な差が出ておらず、予測精度も未だ不十分である。

#### 3.2 変数の寄与率

model Cにおいて、各数値データがどの程度予測に寄与しているかを確認するため、変数の寄与率を用いる。model Cの数値データは入力サイズ6に対し、次の層で1に集約される。この間の重みを $w_i$  ( $i \in [1, 7]$ )とする。各変数の寄与率 $P$ は以下で算出できる：

$$P = \frac{|w_i|}{\sum_{i=1}^7 |w_i|} \quad (3)$$

実際の算出結果は図3である。7日間累積日射量が46%を占め、開花節数が21%、開花節径が14%と続いた。特に7日間累積日射量の寄与が高いことから、日射量も踏まえた特徴量を生成し、予測を行ったことが示唆された。

#### 3.3 Grad-RAMによる解釈

Grad-RAMを用いた視覚的説明は図4である。 $\tilde{L}_{xy}$ の値が大きく、CNNが予測根拠として着目したと考えられる領域ほど赤く、一方で $\tilde{L}_{xy}$ の値が小さく、あまり着目していないと考えられる領域ほど青く描かれている。この結果から、CNNモデルは主に側面の葉に着目していることが分かる。一般に、光合成量と収量の関係性は知られているが、この結果はCNNは光合成が行われる葉を特徴量の1つとして判断したことを示唆している。

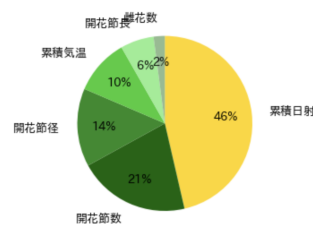


図3: 数値データの寄与率

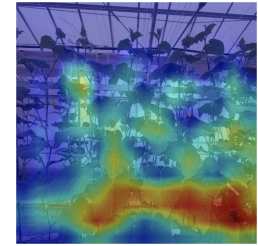


図4: 可視化結果

### 4 まとめ

本研究では1週間後のきゅうりの収量予測をテーマに、画像と数値データを入力とするCNNを用いた回帰を行なった。学習の結果から、画像データのみおよび数値データのみを利用したモデルに比べ、画像データと数値データの両方を使ったモデルは最も精度が向上し、現時点で季節による収量差が分かる程度の精度となった。また、CNNを用いた回帰モデルにおける予測根拠の可視化手法であるRAMを基に、より一般的なCNNモデルに対して適用可能な定義へ変更したGrad-RAMを提案し、本研究のCNNモデルに適用した。その結果や数値データの寄与率の解析から、葉や日射量等を踏まえた特徴量を生成したことが確認できた。

現在は春と秋の各1期のみであることや1区画10枚を利用していることからデータの多様性が不足しているため、今後継続してデータを取得することで精度と頑健性が上がることが期待される。また、今回使用したCNNモデルおよびNNモデルは単純な構造であることから、中間層の数の変更や、画像と数値の情報を結合する位置の最適化をすることで、より精度の高いモデルの構築をすることが今後の課題である。

#### 参考文献

- [1] 濱田敏弘, 山下泰生, 柴田英明: オリーブ果実の収量予測と、オイル用果実一斉収穫推進のための技術開発 (第1報) -オリーブの収量予測手法に関する検討-, 香川県産業技術センター研究報告, 18号, p.60-64, 2018.
- [2] 磯山侑里, 江村文代, 佐藤裕久, 篠崎隆志: 周年栽培トマトの果実検出システムの検証, 第33回人工知能学会全国大会論文集, 2019.
- [3] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva and A. Torralba: Learning deep features for discriminative localization, In *Proceedings of the IEEE conference on computer vision and pattern recognition*, p.2921-2929, 2016.
- [4] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra: Grad-cam: Visual explanations from deep networks via gradient-based localization, In *Proceedings of the IEEE ICCV*, p. 618-626, 2017.
- [5] Z. Wang and J. Yang: Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation, arXiv: 1703.10757, 2017.
- [6] G. Bradski: opencv\_library, Dr. Dobb's Journal of Software Tools, 2000.