

ディープラーニングを用いた植物の生長予測

櫻井 俊輔[†] 内山 英昭[†] 島田 敦士[†] 谷口 倫一郎[†]九州大学[†]

1 はじめに

農業において収穫量の増加や、安定した生育管理は重要な課題である。植物の生長予測はこの課題の解決に貢献する技術の一つである。生長予測の手法として生育環境をパラメータとしたモデル化が考えられるが、生育に影響する環境要因は多様であり、環境条件の厳密な計測には数多くの設備が必要となり高コストである。近年、カメラの低コスト化と画像処理技術の向上が進んでおり、画像処理は生長予測のコスト低下を実現する手法の一つとして期待される。画像処理による手法には植物の非破壊性や、大量かつ高速に適用可能であるという利点が考えられる[1]。このような点から、画像処理による生長予測手法の開発が要求される。

画像処理による予測技術の一つとして、動画の将来のフレームを予測する研究が挙げられる。Srivastavaらは動画のフレーム間の長期的な関係性を学習するために時系列解析に適した構造を持つDeep Neural Network (DNN)であるLong-Short Term Memory(LSTM)を用いている[2]。また、より効果的な動画の内部表現を得るために、タスクを動画の解釈と予測の生成に分割するエンコーダデコーダ構造を適用している。

本稿では、動画のフレーム予測の枠組みを参考に、植物の生長過程を撮影した時系列画像から、今後の生長を示す生長予測画像を生成する手法を提案する。具体的には、生長の時系列性と生長表現の重要性に着目し、エンコーダデコーダ構造のConvLSTMを植物の生長画像に適用する。また、実際の生長を適切に表現する特徴獲得のため二種類の誤差関数を提案する。

2 生長予測ネットワーク

本節では植物の生長予測における提案ネットワークについて述べる。提案手法は過去の画像列から未来の画像を出力とする点で動画のフレーム予測ネットワークに基づいているが、従来

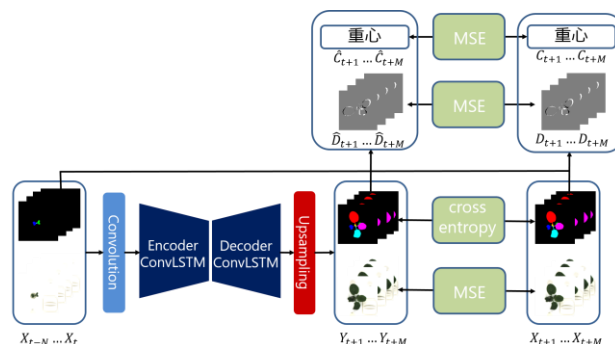


図1: ネットワークの概要図

のタスクとは2.1節に述べる点で異なっている。また、2.2節、2.3節では提案ネットワークの誤差関数について述べる。

2.1 フレーム予測との差異

フレーム予測との差異の一つはRGB画像と葉領域ラベル画像の両方を入出力に用いている点である。植物の表現型解析の観点では葉領域ラベル画像が要求されるが、RGBの予測画像をラベリングする場合にはラベル精度の低下が考えられる。そこで、提案ネットワークではラベル画像も同時に出力する。このため、入力にもラベル画像を利用する。二つ目の差異はフレーム間の実時間である。一般にはフレーム間が一秒未満であるようなごく短い動画が対象となるが、植物の生長にはより長いタイムスケールを必要とする。この仮定のために、文献[3]のようなオプティカルフローに基づく誤差関数の利用は適切でない。そこで、本研究ではフレームに渡る変化に制約を加える差分誤差と葉の動きを直に制約する重心誤差の二種類の誤差を追加する。

2.2 差分誤差

同フレーム内での類似性以外に、他フレームからの変化に関する最適化のために、前フレームとの差分を最小化する差分誤差を提案する。tフレーム目の予測に対する差分画像 \hat{D}_t とその真値 D_t は図1の表記を用いて以下で定義される。

表 1 : KOMATSUNA データセットに対する実験結果の weighted coverage score

	葉 1	葉 2	葉 3	葉 4	葉 5	葉 6	葉 7	平均
提案誤差無し	72.5	68.7	67.8	72.6	66.6	61.1	32.4	67.2
差分誤差	72.1	67.1	67.2	73.4	69.2	60.7	37.5	67.9
重心誤差	71.8	65.2	68.5	72.7	67.7	59.5	34.8	67.1
併用	72.8	64.7	68.2	74.3	69.0	62.2	37.2	68.1

$$D_t = X_t^{label} - X_{t-1}^{label} \quad (1)$$

$$\hat{D}_t = Y_t^{label} - Y_{t-1}^{label} \quad (2)$$

ただし、対応する予測画像が存在しない場合測の代わりに実画像を用いる。差分誤差 L_{diff} は以下で定義される。

$$L_{diff}(X_t, Y_t) = MSE(D_t, \hat{D}_t) \quad (3)$$

2.3 重心誤差

差分誤差に加えて、葉の外見変化に関する誤差として重心誤差を提案する。差分誤差が葉の拡大を最適化する役割を持つ一方で、重心誤差は葉の移動を最適化する役割を持つ。重心誤差 L_{cr} は真値画像から得た葉の重心 C_t と予測画像から得た重心 \hat{C}_t によって以下で定義される。

$$L_{cr}(X_t, Y_t) = MSE(C_t, \hat{C}_t) \quad (4)$$

重心はラベル画像のモーメントから各葉について計算される。

3 実験

提案手法の有効性を検証するため、ネットワークを KOMATSUNA データセット[4]のラベル付けされた植物の生長データに適用した。結果の定量評価指標として予測と真値との weighted coverage score (WCS)を用いた。また、定性評価として予測画像を真値と直接比較した。WCS を計算する際、画像中に存在しない葉は計算から除外した。

表 1 に使用誤差関数ごとの WCS を示す。また、図 2 に予測画像を示す。全ての結果は左から右の順序で配置されている。表 1 より、差分誤差は特に葉の変化が大きくなる後半の葉で精度向上に寄与していることが分かる。重心誤差の追加は単体では一部の葉の予測精度を向上させるにとどまったが、差分誤差と組み合わせることでより高精度での予測を可能とし、平均 WCS は最も高くなった。図 2 の結果においても、重心誤差と差分誤差を組み合わせることで葉の外見がより真値に近い予測が生成されることが見て取れた。

4 おわりに

本稿では、植物の生長予測ネットワークと誤差

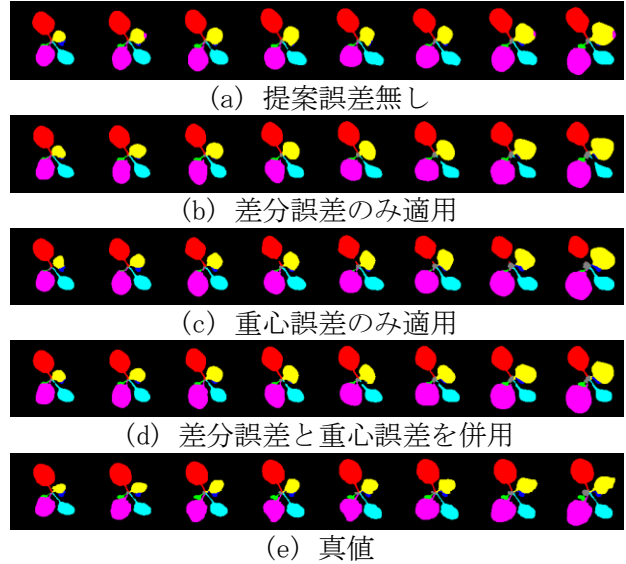


図 2 : ネットワークの生成結果とその真値

関数を提案した。定量的な評価として WCS を、定性的な評価として結果の外観を用い、小松菜の生長予測に対する性能を比較することで、手法の有効性を示した。本稿では入力に真値のラベル画像を用いるという強力な仮定を用いており、RGB 画像のみを入力とした予測は今後の課題である。

参考文献

- [1] Moriyuki, Shogo, and Hirokazu Fukuda. "High-throughput growth prediction for *Lactuca sativa* L. seedlings using chlorophyll fluorescence in a plant factory with artificial lighting." *Frontiers in plant science* 7 (2016): 394.
- [2] Srivastava, Nitish, Elman Mansimov, and Ruslan Salakhudinov. "Unsupervised learning of video representations using lstms." *International conference on machine learning*. 2015.
- [3] Liu, W., Luo, W., Lian, D., and Gao, S. (2018). Future frame prediction for anomaly detection a new baseline. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [4] Uchiyama, Hideaki, et al. "An easy-to-setup 3D phenotyping platform for KOMATSUNA dataset." *Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on*. IEEE, 2017.