

ニューラルネットワークを用いた Q 学習による農作物の栽培に向けた試みCultivation of Crops by Q -learning using Neural Network

難波 脩人[†]
Shuto Namba

辻 順平[‡]
Junpei Tsuji

能登 正人[†]
Masato Noto

1. はじめに

日本における農業問題の一つとして後継者の育成が困難であることが挙げられる。近年の異常気象により、同じ農作物を栽培しているにもかかわらず、毎年作物の状態に合わせた適切な栽培方法が必要とされ、新規就農者にも臨機応変に栽培を行うことが要求される。このような背景のもと、熟練農家の知見を教師あり学習を用いて再現することにより、新規就農者への技術継承を目的とした研究が行われている [1]。一方、我々は熟練農家の知る栽培方法が最適かどうか判断することは難しく、より最適な栽培方法があるのではないかと考え、知見に依存しない強化学習によって農作物の栽培を行うという研究に取り組んできた。これまでにニューラルネットワーク (NN) を用いた Q 学習により、農作物 (小松菜) が状態 (背丈) に合わせて適切に行動 (灌水) を選択している様子を確認してきた。

本研究では NN を用いた Q 学習から得た結果に基づき、小松菜が選択した水の与え方から生育が行われるかを実験的に明らかにする。

2. 強化学習による農作物栽培

強化学習を農作物に適用する際に重要な要素として状態の定義が挙げられる。作物は常に時間変化しているため、1 回の栽培につき、同じ状態を訪れることはなく、 Q テーブルを作成することが困難である。我々はこのような問題に対して、NN を用いて Q 関数を作成する手法を採用してきた。本研究における Q 関数導出の流れを図 1 に示す。本研究では実際に小松菜の栽培を行うための NN の構造として 3 入力と 2 入力の単層構造を採用した。3 入力時における時刻 t での状態 s_t は時刻 $t-2$, $t-1$, t の背丈を表した 1 次元ベクトルを 3 つ並べた 3 次元ベクトルとして定義した。また、2 入力時には同様に時刻 $t-1$, t の背丈を表した 1 次元ベクトルを 2 つ並べた 2 次元ベクトルとして定義した。出力層には tanh 関数を採用することで Q 関数の近似を行っている。行動 (action) は $a^{(0)} = "0 \text{ ml}"$, $a^{(1)} = "10 \text{ ml}"$, $a^{(2)} = "15 \text{ ml}"$, $a^{(3)} = "30 \text{ ml}"$ の灌水の 4 パターンを用意し、状態 s_t で action を起こした場合に次の状態 s_{t+1} に遷移するのは作物への影響が出るための時間を考慮し、試行錯誤した結果から 2 日後に設定している。報

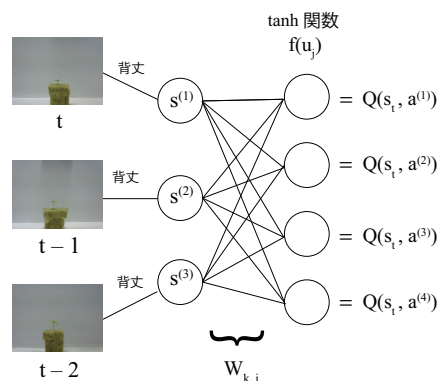


図 1: 3 入力時の NN の構造。2 入力時の構造は時刻 t と $t-1$ の背丈のみを入力としている。

酬の定義は時刻 t と $t-1$ における植物の背丈を比較し、以下のように定義する。

$$r_t = \begin{cases} +0.2 & (\text{時刻 } t \text{ で生長している}) \\ 0 & (\text{変化なし}) \\ -0.2 & (\text{時刻 } t \text{ でしおれている}) \end{cases}$$

上記の条件下で小松菜をエージェントとした強化学習を行う。

強化学習ではエージェントが試行を繰り返すことにより最適解へ収束することができるが、作物をエージェントとした場合、栽培期間の長さから試行を繰り返すことは困難である。そのため、栽培期間は 1 試行を 2 週間とした苗の期間に限定することで強化学習を行う。一方で、1 試行を 2 週間とした場合においても作物の栽培を最適化させるためには膨大な時間が必要になることが考えられる。そのため、学習の効率化を図るために 1 度経験した情報を再利用する experience replay (ER) を用いる。ER はエージェントが経験したサンプル $\{(s_t, a_t, r_t, s_{t+1})\}_{t=1}^T$ をメモリに保存する。このとき、 T はサンプル数の合計を表しており、本研究では $T = 256$ として学習を行う。農作物に強化学習を適用した研究としてシミュレーションを用いて収穫量の最適化を行う研究 [2] があるが、実環境で農作物を栽培することから手法の評価を行う研究は十分に議論されていない。本研究では実際に作物の栽培を通して人の手によって栽培した小松菜の背丈と強化学習によって栽培した小松菜の背丈を比較することで NN

[†] 神奈川大学, Kanagawa University

[‡] 東洋大学, Toyo University

表 1: 小松菜の栽培結果

	植えた個体数	発芽した個体数	生長した個体数	枯れた個体数	最終背丈	
					平均 (cm)	標準偏差
人の手	16	16	16	0	1.54	0.39
2 入力	40	39	39	0	1.49	0.49
3 入力	40	40	24	16	0.7	0.6

を用いた Q 学習から農作物の栽培が可能かどうか評価する。

3. 結果と考察

NN を用いた Q 学習から得た結果に基づき、小松菜が選択した水の与え方から生育が行われるかを明らかにするため、2019 年 5 月 4 日から 2019 年 5 月 17 日までの 2 週間で 96 株を対象に実験を行った。本研究では人の手によって栽培した小松菜が 16 株、強化学習を用いて栽培した小松菜が 40 株ずつの計 96 株の栽培を行った。

表 1 に上記の期間による小松菜の栽培結果を示す。人の手による栽培と 2 入力の NN による栽培は発芽した個体すべての生長を確認することはできたが、3 入力の NN による栽培は 24 株が生長し、16 株が枯れる結果となった。図 2 のグラフからも 3 入力で学習した NN から得られる行動は人が選択している行動とは異なるため、枯れる個体が現れたことが考えられる。一方で、表 1 の最終背丈の平均値からは 2 入力型の NN による栽培は人の手による栽培結果と比較しても有意差があるとは言えないことがわかる。小松菜の生長曲線 (図 3) を見ても人の手による栽培結果と 2 入力による栽培結果を比較して人が育てた場合と近い生長曲線を描いていることがわかる。一般的な Q 学習では困難であった作物の栽培に強化学習を適用する際の手法として NN を用いた手法が有効であることを示唆している。

4. おわりに

本研究では NN を用いた Q 学習によって農作物の栽培が可能かどうか評価することを目的とし、実際に小松菜の栽培を行うことによって実験的に明らかにした。結果として、NN の構造を状態の入力数を 2 次元ベクトルで表した構造において小松菜の最終背丈が人による栽培結果と比較して有意差があるとは言えないことが明らかになった。しかしながら、人による栽培結果と比較した際に最終背丈の平均でわずかに劣ることからも今回の強化学習の結果が局所最適解に収束している可能性が考えられる。局所最適解に収束している要因の一つとして報酬の設定が疎な報酬になっていたことが考えられる。人による栽培を行う際には背丈の情報のみではなく栽培に使用しているウールに含まれる水分量などを考慮した栽培を行っていたが、強化学習による栽培では報酬に背丈の情報しか与えておらず、学習がうまくいかなかったことが考えられる。人の手によって報酬を設計するので

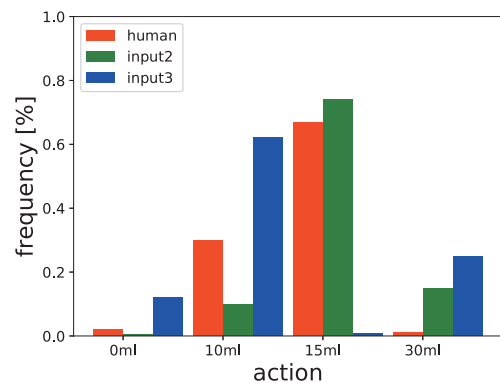


図 2: action の出現頻度

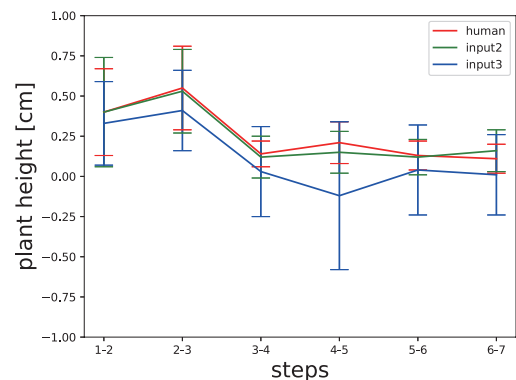


図 3: 小松菜の生長曲線。各ステップ間での生長幅の平均と標準偏差を表している。

はなく逆強化学習などの手法を用いて報酬関数の推定を行い手法の改善を進めたい。

参考文献

- [1] Mohanty, S. P., Hughes, D. P. and Salathé, M.: Using Deep Learning for Image-based Plant Disease Detection, *Frontiers in Plant Science*, Vol. 7, pp. 1–10 (2016).
- [2] Lijia, S., Yanxiang, Y., Jiang, H., Dana, P., Thomas, M. and Charles, H.: Reinforcement Learning Control for Water-Efficient Agricultural Irrigation, pp. 1334–1341 (2017).