

令和 5 年 6 月 10 日現在

機関番号：14301

研究種目：基盤研究(B)（一般）

研究期間：2020～2022

課題番号：20H02968

研究課題名（和文）深層学習に基づくイネバイオマスの汎用的推定モデル構築とその応用

研究課題名（英文）Deep learning-based estimation of rice biomass and its application

研究代表者

田中 佑（Tanaka, Yu）

京都大学・農学研究科・助教

研究者番号：50634474

交付決定額（研究期間全体）：（直接経費） 13,500,000円

研究成果の概要（和文）：作物の生育を正確に把握することは、生産現場における栽培管理や、育種現場における高生産性系統の選抜など、多様な場面で必須である。本研究では重要作物であるイネを対象とし、深層学習を活用することで、可視画像のみから簡便かつ正確にバイオマスを推定するモデルの開発を目指した。多地点で栽培された多様なイネ群落の画像と対応するバイオマスの実測値から深層学習モデルを構築した。得られたモデルは、未知の地点で栽培されたイネ品種に対しても十分な精度でバイオマスを推定できることが示され、省力的かつ非破壊でのバイオマス推定技術を構築することができた。

研究成果の学術的意義や社会的意義

作物の安定多収化を実現するためには、作物のバイオマスを正確かつ簡便に把握することが極めて重要である。本研究で得られた技術は、市販のデジタルカメラやスマートフォンを利用することで、イネの生育量を即座に推定できることを示している。生産現場において、本技術を活用することで、イネの適切な施肥管理や圃場管理が容易になると期待される。さらに世界的には、イネのさらなる多収化が重要な課題であり、育種現場において、大量の遺伝資源から有望系統を選抜するうえでも有効なツールになると考えられる。

研究成果の概要（英文）：Monitoring the crop growth is the fundamental process for all the aspects of the crop science. There have been, however, few practical methods enabling the non-invasive, efficient and accurate estimation of the crop biomass. Here, we propose to apply the deep learning technology to develop the model for estimation of rice (*Oryza sativa* L.) biomass based on the RGB image.

The database of the RGB images of rice canopy and corresponding biomass covering wide range of growth environments and cultivars was established. The deep learning model developed based on the database showed the high accuracy to estimate the rice biomass. Notably, the accuracy was practical even when the target cultivar and growth environment was not included in the database used for the model development.

Our study demonstrated the efficient and accurate estimation of rice biomass only by using RGB images, which is expected to accelerate the analysis of crop growth under various situations.

研究分野：作物学

キーワード：イネ 深層学習 バイオマス 可視画像

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

本研究が対象とするイネにおいて、近現代の育種による収量向上は主に収穫指数(HI)の向上により達成されており、HIには改良の余地が少ないとされる。したがってさらなるイネの生産性向上を目指す場合、面積あたりの地上部全乾物重(バイオマス)の増大が必要である。一方で、生産基盤が確立されていない発展途上国の生産現場では、農家の栽培技術、作付様式、品種、土壌肥沃度、水利条件、気象など時空間的に様々な生産規定要因が混在するために、隣接する圃場間でも収量やバイオマス生産量が大きく異なる。こうした地域では、多様な生産圃場におけるバイオマス生産量の実態を迅速かつ広域に把握し、その制限要因を解明することで、圃場毎に最適化された効果的な生産性改善策に繋げることが重要となる。以上のように、イネのバイオマス測定は育種・栽培学の両面において根幹をなす要素である。

研究開始当初におけるバイオマスの測定は、直接的な刈り取り調査が基本である。これは一定個体数/面積の植物体の地上部を人力で刈取り、乾燥重を測定する手法である。したがってバイオマス生産過程を継続的に調査する場合、大規模な刈取り用プロットをあらかじめ準備する必要がある。しかも直接的な刈取りには大きな労力がかかるほか、生産圃場においてはそもそも継続的な刈取り調査が困難なケースが多い。代替手段として、スペクトル情報を人工衛星や無人航空機(UAV)から測定し、バイオマスを非破壊推定する手法などが多数提案されている。しかしこれらの推定結果はイネ品種の草型などに大きく影響を受けるため、対象地点におけるバイオマスの実測値無しでは精度が担保されない。そのため汎用性が高いとはいえ、高度な機材と知識を必要とすることから、バイオマスの直接測定にかわる手法とはなっていない。すなわち【必要な時と場所で】、【高価な専用機材に依存せず】、【十分な精度を以て】、【迅速に】バイオマスを非破壊に推定するための汎用的方法は未確立であり、このことが作物のバイオマス研究において根本的なボトルネックとなっていた。

2. 研究の目的

本研究の最終目的は、簡易・高精度かつ汎用的なイネのバイオマス非破壊推定システムを構築し、多収選抜育種や多様な栽培環境にある生産圃場での生育量把握に貢献する技術基盤を構築することである。そのため、以下の3点を目標とした。

- (1) 近年著しく進歩しつつある深層学習の手法を応用し、イネの群落可視画像から多様な草型に対応した汎用的なバイオマス推定モデルを作成すること。
- (2) 国際イネ研究所(以下 IRRI)が所有する大規模集団を対象として高バイオマス系統の選抜とそれを支配する遺伝的要因の解析を行うこと。
- (3) マダガスカルが生産圃場においてバイオマスの非破壊推定を行い、広域にわたる生育量把握とそれに基づく収量変動要因解析を行うこと。

3. 研究の方法

(1) イネ群落画像-バイオマスデータベースの構築

(i) 2株単位のモデル

深層学習では、入力画像と実測値(バイオマス)のデータベースを構築することが必要となる。モデルの汎用性と性能は、データベースを構成する画像やバイオマスの多様性に左右される。本研究ではまず、イネの遺伝的多様性を最大限網羅した世界のイネコアコレクションの中から59品種を対象とし、京都、および東京の2地点で2年にわたり栽培した。2579の刈取りプロットを設定し、2株単位のイネ画像を取得した直後、刈り取りによりバイオマスデータを得た。画像は市販のデジタルカメラを用いて地上150cmから鉛直下向き、焦点距離28mm(35mmフィルム換算)で取得し、上下左右反転処理などにより、合計12183枚得られた。

(ii) 1m²単位のモデル

世界のイネコアコレクションをはじめとする多様なイネ品種群を、京都、千葉、マダガスカル、コートジボアールの4地点で栽培した。1911の刈取りプロットを設定し、1m²株単位のイネ画像を取得した直後、刈り取りによりバイオマスデータを得た。画像の撮影方法については、群落表面から80-90cmの高さから鉛直下向きに取得し、その他の設定は(i)と同様とした。

(2) データの分割とモデルの構築

前項いずれのデータベースにおいても、モデルの学習に用いる training、モデル選択に用いる validation、精度検証に用いる test および完全に独立したデータセットとしての prediction に大別し、解析に供した。本研究では、深層学習の中心的計算手法である畳み込みニューラルネットワーク(CNN)を構築した。CNNでは、計算ネットワークを自由に設計することが出来、それがモデルの性能に影響することが知られている。そこでまず、Neural Network Console version 1.5 (Sony Network Communications)を用いて、2株単位のデータベースを入力し、自動構造探索機能を用いて、最も誤差が小さくなるCNNの構造を探索した。結果得られた構造をNNC6と名付け、プログラミング言語Python上で深層学習ライブラリTensorFlow / Kerasにより実装し、

以下全ての解析に用いた。

4. 研究成果

(1) 2株単位のモデル

CNN モデルの構造探索の結果、総パラメーター数 41017 からなる構造が最適として選択された。CNN モデルによる学習の結果、コアコレクションのうち training データにも含まれていた 28 品種から構成される test データに対するバイオマス推定精度は $R^2=0.95$ 、相対 RMSE=18.6% となり、高い精度が得られた。生育ステージとしては出穂後 1 週間頃まで高い推定精度を維持した。学習に用いられたなかった 31 品種から構成される prediction データに対しても、全体として大小の対応関係はみられたものの、実測バイオマスが高い場合にモデルが過小評価する傾向がみられた。折れ線回帰を実施したところ、924 gm⁻² に限界点があると解析され、それ以下のバイオマスに対しては $R^2=0.87$ 、相対 RMSE=27.7% の推定精度が得られた。一方、それ以上のバイオマスに対しては $R^2=0.02$ 、相対 RMSE=34.2% となり、有効な推定を行うことが出来なかった。

作物の生産性関連形質の予測や評価においては、データ取得の労力の面から、同一の実験系で得られたデータセットを training 用と validation 用に分割し、validation 用でモデルの精度評価を行うことが多い。本研究ではこれに加え、training データとは独立した prediction データに対する精度評価を行ったことが特色である。prediction データに対しても、バイオマスが 924 gm⁻² 以下であれば、実用的な精度でバイオマスを推定することが出来た点は、本手法の応用可能性を示唆するものである。さらに、対象とした世界のイネコアコレクションは、草丈や葉色など形態的特徴が極めて多様性に富んでおり、それらを包含し高い推定精度を示した本手法は、様々なイネ品種に対して高い汎用性を有していると期待される。一方で、924 gm⁻² 以上のバイオマスに対して推定が困難であった理由については、

- ・同領域はかなりの高バイオマスに相当し、そのような training データが不足している

- ・葉群が完全に群落を覆うと、バイオマスの大小を評価するための情報が得られにくくなるなどが考えられた。さらに 2 株単位のモデルでは推定領域が狭いため、圃場条件でのイネ生産性評価には実用性に欠けることが明らかとなった。そこで当初計画にはなかった、撮影・推定領域の拡大を行った。

(2) 1m² 単位のモデル

(1)と同様の手法により、1m² のイネ群落画像と実測バイオマスからなるデータベースを構築し、CNN モデルを用いて学習を行った。その結果、training データにも含まれていた品種から構成される test データに対して、 $R^2=0.87$ 、相対 RMSE=0.30 の精度でバイオマス推定が可能であった。一方、training データとは異なる地点で栽培された異なる品種で構成される prediction データに対しては、 $R^2=0.52$ 、rRMSE=0.59 と大きく推定精度が低下した。これは、撮影・推定領域を 2 株から 1m² に拡大した結果、画像の複雑性が増すと同時に、労力の制約から実測のバイオマスデータ数が限られたため、モデルの推定力が不足した結果であると考えられた。

現実的な労力の範囲内で、イネ画像と対応したバイオマスデータの量を増強するための対策を考案した。まずはイネ群落の画像のみを撮影し、中央の 2 株のみをトリミングする。次いで、構築済の 2 株モデルを用いてバイオマスを推定し、得られた推定値を 1m² 画像全体に対する真値として扱った。すなわち、2 株モデルの精度が十分に高いこと、および 1m² のうち中央 2 株のバイオマスがおよそ全体を代表することを仮定した。これにより、あくまでも疑似ラベルではあるものの、画像取得のみでデータ量を増強することにより、データベースの規模を 2 倍に拡張した。拡張したデータベースで学習を行った CNN では、test データに対する推定精度が $R^2=0.91$ 、相対 RMSE=0.26、prediction データに対する推定精度が $R^2=0.79$ 、rRMSE=0.48 と、大幅に改善した。2 株モデルの推定値を真値として扱うことに由来する誤差よりも、データ量の拡張によるメリットが大きかったため、CNN モデルの精度向上に成功したと解釈された。

得られたバイオマス推定モデルは、非破壊で多数のイネ系統の生育を評価するうえで強力なツールになると考えられる。そこで、研究計画にあった国際イネ研究所 (IRRI; フィリピン) での多様なイネ遺伝資源コレクションのバイオマス非破壊評価を行った。具体的には 2023 年 1 月下旬に IRRI に滞在し、出穂前後の 652 品種を対象として、市販のデジタルカメラを用いて 1m² の範囲のバイオマスを推定した。このうち、視野内に欠株が無く、かつ出穂前であった 317 品種のバイオマスには、100 gm⁻² 以下から 1000 gm⁻² 以上にわたる極めて大きな遺伝的変異が見出された。同解析は一人が一日で終わられる作業量であり、従来の刈取りによるバイオマス実測と比較し本手法の有用性が実証された。

(3) 結論

本研究により、深層学習に基づきバイオマスを非破壊かつ簡便に、可視画像から推定する手法が開発された。推定精度についても、分光反射特性などを利用した既存のリモートセンシング技術と比較して遜色ない技術となった。病虫害、倒伏など栽培現場における多様な状況への対応は今後の課題であるものの、その簡便性や汎用性は既存の手法を凌駕するといえる。新型コロナウイルスの拡大により、当初予定していた海外調査を大幅に縮小せざるを得なかったものの、国際イネ研究所 (IRRI) における解析は今後の本手法の応用可能性を示唆するものであり、UAV などによる効率的な画像取得方法と組み合わせることで、これまで不可能であったイネ生育量の簡

便な面的把握，および多数の系統を対象とした高バイオマス品種の効率的選抜が実現すると期待される．

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Nakajima Kota, Tanaka Yu, Katsura Keisuke, Yamaguchi Tomoaki, Watanabe Tomoya, Shiraiwa Tatsuhiko	4. 巻 2
2. 論文標題 Biomass estimation of World rice (<i>Oryza sativa</i> L.) core collection based on the convolutional neural network and digital images of canopy	5. 発行年 2023年
3. 雑誌名 Plant Production Science	6. 最初と最後の頁 1~10
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/1343943X.2023.2210767	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計1件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 中島 洸太, 田中 佑, 桂 圭佑, 山口 友亮, 白岩 立彦
2. 発表標題 多様な草型を有するイネ品種を対象とした深層学習モデルによるバイオマス推定
3. 学会等名 日本作物学会第251回講演会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	桂 圭佑 (Katsura Keisuke) (20432338)	東京農工大学・(連合)農学研究科(研究院)・准教授 (12605)	
研究分担者	辻本 泰弘 (Tsujimoto Yasuhiro) (20588511)	国立研究開発法人国際農林水産業研究センター・生産環境・畜産領域・プロジェクトリーダー (82104)	
研究分担者	高井 俊之 (Takai Toshiyuki) (40547725)	国立研究開発法人国際農林水産業研究センター・生産環境・畜産領域・主任研究員 (82104)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------