Rice yield prediction using UAV-based multispectral imagery and weather data through multimodal deep learning

○Md Suruj Mia (UGSAS, Gifu Univ.), Ryoya Tanabe (Grad. Sch. Nat. Sci. Technol., Gifu Univ.), Luthfan Nur Habibi (UGSAS, Gifu Univ.), Naoyuki Hashimoto (Fac. Agric. Marine Sci., Kochi Univ.), Koki Homma (Grad. Sch. Agric. Sci., Tohoku Univ.), Masayasu Maki (Fac. Food Agric. Sci., Fukushima Univ.), Tsutomu Matsui and Takashi S. T. Tanaka (Fac. Appl. Biol. Sci., Gifu Univ.)

1. Introduction

Timely, nondestructive, inexpensive, and precise large-scale yield forecast is essential for preventing climate risks and ensuring food security. Satellite-based remote sensing for crop yield prediction has often been constrained by the limited spatial and temporal resolution of satellite data. In response to the challenges, unmanned aerial vehicles (UAVs) have been widely used to collect data due to their superior spatial and temporal resolutions and flexibility in sensor selection compared to airborne and satellite platforms.

Machine learning algorithms possess the capability to devise novel solutions for real-world challenges, enabling farmers to make informed decisions while requiring minimal or even no human intervention. The Convolutional Neural Network (CNN) algorithm is efficient not only in image classification tasks but also for regression tasks such as crop yield prediction. Weather is one of the major environmental factors affecting crop growth and yield. Integrating weather and image data as input layers into CNN models may enhance the accuracy of the yield prediction. The objective of this study was to develop a multimodal deep learning model to predict rice grain yield using UAV images at the heading stage and aggregated weather data.

2. Materials and Method

Rice yield surveys were conducted in 22 farmers' fields in Japan over six years (2017–2022). Nine rice varieties were planted, and crop management was conducted according to local conventional methods. Strip trials were performed in 12 fields for basal fertilizer application to obtain high yield variability and to determine whether the effect of fertilizer application rate on rice yield could be evaluated using a predicted yield map. A total of 894 samples were collected in all yield surveys, from an area of approximately 1.0 m^2 at the heading stage. UAV multispectral images and weather data (https://amu.rd.naro.go.jp) for each region are collected and processed. 18 architecture models consisting of two types of CNN feature extractor layers (i.e., CNN_2conv and AlexNet), three different depths of fully connected layers, and three methods of integrating weather data (no weather data, weekly sum of weather data, and monthly sum of weather data) into deep neural network models were established. The model performance was measured using the R^2, root mean squared error (RMSE), and root mean squared percentage error (RMSPE) values.

3. Results and Discussion

The collected yield data exhibited a normal distribution. There was no significant difference in model accuracy between the AlexNet and CNN_2conv architectures. The effects of the number of layers and weather data types on the model performance were significant but according to Tukey's HSD results, there were no significant differences among various layers. Models trained with weekly weather data consistently outperform others. Thus, the simpler CNN extractor (i.e., CNN_2conv) and weekly aggregated weather data might be the best combination for multimodal deep learning in terms of yield prediction accuracy and computational cost. However, spatial distributions of predicted yield were inconsistent between the best two models, although the models showed almost the same prediction accuracy.