Precision agriculture classification using convolutional neural networks for paddy growth level

Indri Neforawati¹, Nanna Suryana Herman², Othman Mohd²

¹Politeknik Negeri Jakarta, Depok, Indonesia
²Universiti Teknikal Malaysia Melaka, Malaysia

Corresponding Author Email: indri.neforawati@tik.pnj.ac.id

Abstract. Precision Agricultural is a key component of modern agricultural. Several researchers tried to use various machine learning models as precision agricultural classification and recognition model, but surprisingly merely few researchers use Deep learning models to solve precision agriculture problems like Paddy Classification, Plant Classification or Fruit Classification. In this research, Precision Agriculture Classification on Paddy Image Dataset was performed using Convolutional Neural Networks. Paddy should be catered well in order to monitor time to harvest, time to watering, and other tasks. The result of classification, we obtained 82% overall accuracy.

1. Introduction

Rice or paddy is an essential primary food for human especially in South East Asia such as Indonesia, Thailand, Malaysia, and others. The population of paddy should be catered in order to keep sustainability of primary food for human. Monitoring and prediction of paddy growth can be taken by technology like Precision Agriculture in order to enhance farming techniques. Precision Agriculture involves Geographic Information System (GIS), Global Positioning System (GPS), Remote Sensing, and classification techniques. GPS usually used to determine the location of the object, Remote Sensing usually used as Data Gathering, meanwhile, GIS used to show the result of Precision Agriculture. There are several Precision Agriculture classification techniques that can be used like Machine Learning or Deep Learning. Previous research by Sadiyah et al [1] produces 99.58% overall accuracy of plant image dataset using Convolutional Neural Networks, meanwhile other Sadiyah et al paper [2] produce 99.8% overall accuracy using the same model. The using Deep Learning model as classification techniques to Precision Agriculture dataset can produce positive performance comparing to others, so in this paper, we try to use the same model to classify other Precision Agriculture dataset: paddy. Paddy should be monitored well like time to harvest, time to watering and Growth Level of Paddy. The last afore mention will be the priority of this research. By knowing the growth level of paddy, farmers can determine when paddy should be harvest and water. The contributions of this paper are: (1) to try Convolutional Neural Networks model to Paddy Dataset, (2) to determine the growth level of paddy.
2. Methods
The methods starting from data collection of Paddy Dataset. We used the Food and Agriculture Organization of the United Nations (FAOSTAT) dataset which can be used as the open source. The dataset is also enriched with other datasets particularly by our own dataset. After collecting data, the dataset will be pre-processed to enhance image quality and reduce alpha channel in order to make data can be readable in the next step. Dataset will be separated into three parts: Train, Validation, and Testing data. Classification techniques Convolutional Neural Networks (CNN) [3] in figure 1 will be performed to create the model using training data that can determine the growth level of paddy. After having the appropriate model, the evaluation will be performed using Precision, Recall, and F-Score. Testing data is also feed-forwarded into final model to obtain the prediction of growth level of paddy.

![Figure 1. Convolutional Neural Networks](image1)

Convolutional Neural Networks model can receive 3-channel images (Red, Green, and Blue) into the input layer, then convolutional layer \( \text{Conv}(x, W) \) will be performed to obtain features automatically. The output from the Convolutional layer will be feed-forwarded into Sub-Sampling layer \( \text{Sampling}(x, W) \) which also obtain feature automatically. This process will be repeated until meet fully connected layer which converts the layer into the one-dimensional vector. The final step, in the classification layer \( \text{Softmax}(z_i) \) will be performed to determine which class has better score result.

\[
\text{Conv}(x, W) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} W_{(i-m,j-n)}
\]

\[
\text{Sampling}(x, W) = \sum_{i=1}^{n} \sum_{j=1}^{m} \text{Conv}_{ij} W_{(i-m,j-n)}
\]

\[
\text{Softmax}(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^{n} \exp(z_j)}
\]

Other deep learning techniques like Batch Normalization [4], Dropout [5], and Adam Optimization techniques [6] will be also used in this architecture and model in order to improve performance.

![Figure 2. Paddy Dataset](image2)
In order to make the comparison between CNN and other methods to classify paddy growth level, we use other Machine Learning model, that is Support Vector Machine (SVM) as a benchmark model. SVM perform well when the data characteristics are linearly separable. In order to use SVM, we use manually feature extraction using NDVI and other techniques.

3. Result
After performing 10 epochs for training data, loss function was reduced as a time reversely accuracy also increase as a time. Training process involves training and validation data as can be seen in figure 3.

![Figure 3. Train and Loss of CNN](image)

We produce overall accuracy 82% of classifying paddy growth level using CNN, on the other side when performing SVM, we merely produce 25% of accuracy. For Precision (P), Recall (R), and F-score (F) measurement evaluation, the experiment model also produce the positive result as can be seen in the below table:

|                  | Precision | Recall | F1 Score | Accuracy |
|------------------|-----------|--------|----------|----------|
| Overall CNN      | 0.82      | 0.81   | 0.78     | 0.82     |
| Overall SVM (Benchmark) | 0.23      | 0.27   | 0.29     | 0.22     |

4. Discussion
We separated the dataset into four classes: Seeding, Transplanting, Flowering, and Maturity. The process of labelling data was done manually. Convolutional Neural Networks (CNN) can produce the better result of Paddy Growth level classification comparing to SVM it causes by the ability of CNN to extract the feature of paddy automatically that different from manually feature extraction by SVM. We produced overall accuracy 82% that means when testing 100 unseen testing data into the model. We obtained 82 correctly prediction testing and remaining produce incorrectly prediction. We also measured the performance of the model using Precision, Recall, and F1 Score of CNN model which produced better result comparing with SVM model.

5. Conclusion
We used current paddy dataset from the Food and Agriculture Organization of the United Nations (FAOSTAT) dataset that augmented by our own dataset. Precision Agriculture classification using Convolutional Neural Networks can produce 82% accuracy better than SVM model. Measurement evaluation using Precision, Recall, and F-score are also producing the better result with 0.82, 0.81, and 0.78 respectively. When mapping testing data into ground truth, we produced positive result that
indicated by Confusion Matrix. For future research, we plan to augment the precision agriculture dataset with the variety of paddy growth level. We also plan to use other deep learning models beyond CNN.

6. References

[1] Abdullahi H S, Sheriff R E and Mahieddine F 2017 Convolution neural network in precision agriculture for plant image recognition and classification 2017 Seventh Int. Conf. Innov. Comput. Technol. pp 1–3.

[2] Sadiyah A Hand Mustpha Z O 2017 Advances of image processing in Precision Agriculture: Using deep learning convolution neural network for soil nutrient classification.

[3] Krizhevsky A, Sutskever I, and Geoffrey E H 2012 ImageNet Classification with Deep Convolutional Neural Networks Adv. Neural Inf. Process. Syst. 25 pp. 1–9.

[4] Wei Y 2015 Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift 2015 Int. Conf. Mach. Learn. vol. 37 pp. 448–56.

[5] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 Dropout: A Simple Way to Prevent Neural Networks from Overfitting J. Mach. Learn. Res. vol. 15 pp. 1929–58.

[6] Ruder S 2016 An overview of gradient descent optimization algorithms.