The convolutional neural network (CNN) is a neural network that is composed of one or more convolutional and pooling layers in the architecture. Array data such as images, local groups of values are often highly correlated, forming locally distinctive features. The convolutional layer can detect such spatial features while the pooling layer can merge semantically similar features into one. The CNN has been originally applied in speech recognition, document reading, and object detection in images (LeCun et al., 2015). Because of the higher performance of the CNN in feature extraction and prediction than classical approaches, the CNN are nowadays applied in agricultural studies such as image-based plant disease detection (Sharma et al., 2020) and crop yield estimation (﻿Nevavuori et al., 2019). Furthermore, a recent study demonstrated that the CNN model had a higher prediction accuracy than classical machine learning algorithm such as random forest in modeling crop yield response to management and environmental variables because the CNN was expected to capture spatial structure of different attributes (e.g. elevation and soil map) affecting crop yield (Barbosa et al., 2020).

In this study, one of the multi-stream CNN architectures proposed by Barbosa et al. (2020) (called Late Fusion), was used with modifications. Briefly, each input size is 6 × 6 (e.g. alpha, beta, etc.) excluding the nitrogen rate. Although the spatial resolution of input size is finer in the CNN model than the other machine learning models (e.g. random forest, causal forest, etc.), the area of each cell (X × X m) is the same among the models. On the other hand, the input size of the nitrogen rate is treated as 1 × 1 because it is spatially homogenous within each cell. Firstly, each input is connected to an independent convolutional layer with eight 3 × 3 filters each with stride one, followed by a 2 × 2 max-pooling layer with stride two. Then, a fully-connected ﻿rectified linear unit (ReLU) layer with sixteen neurons is added after each max-pooling layer, followed by a single ReLU neuron. Finally, multiple neurons are concatenated and fed to the fully-connected ReLU layer with sixteen neurons, followed by an output with a linear activation function. Barbosa et al. (2020) demonstrated that this architecture had the best performance for modeling crop yield response to nitrogen rate management among several CNN architectures. The CNN model was implemented in Python v3.7.6 using Pytorch v1.7.0 (Paszke et al. 2017). The Adam optimizer (Kingma and Ba, 2014) was used with a learning rate of 0.001% (default value). To avoid overfitting, early stopping was used monitoring validation loss with a patience of 10 epochs.

References

﻿Barbosa, A., Trevisan, R., Hovakimyan, N., & Martin, N. F. (2020). Modeling yield response to crop management using convolutional neural networks. Computers and Electronics in Agriculture, 170, 105197. https://doi.org/10.1016/j.compag.2019.105197

﻿Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

﻿Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature 521 (7553), 436–444.

﻿Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture, 163, 104859. https://doi.org/10.1016/j.compag.2019.104859

﻿Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., & Lerer, A. (2017). Automatic differentiation in PyTorch. In: NIPS-W. <https://openreview.net/forum?id=BJJsrmfCZ> (Last accessed 21/7/6)

Sharma, P., Berwal, Y. P. S., & Ghai, W. (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. *Information Processing in Agriculture*, *7*(4), 566–574. https://doi.org/10.1016/j.inpa.2019.11.001