Identification of a Machine Learning Architecture for Potato Disease Classification Using Leaf Images

M.H. Gayantha[1*], W.M.C.J.T. Kithulwatta[2], and R.M.K.T. Rathnayaka[1]

[1]Department of Physical Sciences and Technology, Faculty of Applied Sciences,
Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka.
[2]Faculty of Technological Studies, Uva Wellassa University of Sri Lanka, Badulla, Sri Lanka.

[*]hashan.gayantha1@gmail.com

Agriculture is a vital economic sector for any country. Therefore, improving the agriculture sector properly is much needed; especially computer technology has been widely used for improving the productivity of the cultivation. Potato is one of the mostly consumed vegetables, but it is susceptible to a variety of diseases. The commonly used way to identify them is by looking at the leaves of the plant. Therefore, if artificial intelligence can be utilized to recognize them, it can be used to provide immediate responses. Accordingly, the main aim of this study is to identify a machine learning algorithm that is fast, and more accurate to identify diseases that affect the potato leaf such as late blight and early blight, using multiclass image classification. Image data for the study was acquired from the Potato disease library of the Kaggle repository. It contains three subclasses named early blight, healthy, and late blight which contains 2152 images in the above three classes. VGG16, efficientNetB0, and Support Vector Machine (SVM), Extreme Gradient Boost (XGBoost) were chosen as deep learning-based algorithms and machine learning algorithms, respectively. Initially, models were trained and a test split was done with Keras preprocessing library. Then a comparison was done among the classifiers, considering the accuracy, validation, and loss, and then the best one was selected for the potato disease identification. The factors and parameters affected to increase the power of the models were then considered. The test accuracies achieved by the classifiers were 92%, 95% for CNN models, 83% for SVM, and 86% for XGBoost as average model accuracy approximately in testing phase. These findings would lead to the development of a model that is best suited for detecting potato leaf diseases. Future researchers will be able to program a flying drone using the aforementioned model and computer vision to identify plant diseases immediately.

Keywords: Computer Vision, Deep Learning, EfficientNet, Machine Learning, Support Vector Machine, VGG16

CS-ML-09