Deep Learning for Anthracnose Diagnosis in Turnip Leaves

Tarun Suresh, Mission San Jose High School

Abstract

Fungal diseases in plants are extremely pressing issues in the agricultural industry, threatening global food security by reducing crop yields and quality. Traditional approaches to disease diagnosis and management have failed to recognize symptoms when they first appear. Leaves of the turnip, a plant of high agricultural value, has been especially affected by the fungal disease Anthracnose. Therefore, this study aimed to develop a novel convolutional neural network that can identify turnip leaves with early symptoms of Anathrosce blight. The model had 4 convolutional blocks and was trained on a custom dataset of 1,470 images, randomly split into 60% train, 20% validation, and 20% test. To compare how the CNN model fared with other machine learning algorithms, a support vector machine(SVM) model was developed and trained with the same image dataset. The CNN model's accuracy 98.75% compared to the SVM model's 80.50% accuracy. These results validate the efficacy of the CNN model to accurately identify infected turnip leaves and demonstrate that it can be implemented into a practical disease diagnosis system. Future studies are warranted to improve the model through means such as k-fold cross validation as well as apply the model architecture to other crops and diseases.

Introduction/Objective

- Fungal diseases in plants remain a huge issue in the agricultural industry to this day.
- Conventional approaches to fungal disease management involving diagnosis by visual inspection have largely failed to identify the early symptoms of fungal blight in plant leaves until the disease becomes widespread and severe, significantly reducing crop yield and quality.
- With the advent of novel machine learning, computer vision, and artificial intelligence approaches and plant image datasets, progress has been made developing automated models that can accurately identify fungal disease in various plant species.
- Although significant progress has been in using computer vision to diagnose disease in plants, many plant species of significant agricultural value have yet to be investigated.
- Therefore, this work proposes a convolutional neural network(CNN) that can classify healthy turnip leaves and those infected by the fungal disease Anthracnose, which has been difficult to diagnose in turnip leaves in its early stages.

Figure 1: LeNet architecture of a sample convolutional neural network.

Materials and Methods

Image Data Acquisition and Preprocessing

- A total of 1470 images of healthy and Anthracnose-infected turnip leaves was acquired.
- The images were randomly split into train, test, and validation groups reflecting 60%, 20%, and 20% of the dataset.

Convolutional Neural Network

- 4 convolutional blocks each with "Conv2D" and "MaxPooling2d" layers were implemented in the CNN model.
- Dense layers were added that make binary predictions with a sigmoid activation function.
- The callbacks EarlyStopping and ModelCheckpoint were implemented to save the best model with lowest validation loss for future evaluation.
- ADAM optimizer and binary cross-entropy were used as optimizer and loss function, respectively.
- Data generator was used to draw images from the dataset in random batches, and the model was fit with 30 epochs.
- The training and validation data accuracy and loss are summarized in Figure 4 and Figure 5.
- The accuracy of the best CNN model saved during the training process was evaluated using the testing data.

Support Vector Machine

- Image features were preprocessed, and linearly transformed using libraries such as HOG feature vector and PCA.
- The SVM model was then trained and scored(Figure 6).
- An AUC-ROC curve shown in Figure 7 was used to evaluate model specificity and sensitivity.

Figure 2: Matplotlib visualization of sample diseased turnip leaf(left) and healthy turnip leaf(right) image from training dataset.

Figure 3: Architecture of convolutional neural network

Results

Figure 4: Training(red) and validation(green) accuracy of CNN model versus number of epochs.The training accuracy and validation accuracy of CNN model reached 97.22% and 98.53%, respectively after 6 epochs.

Figure 5: Training(red) and validation(green) loss of CNN model versus number of epochs.The training accuracy and validation loss of CNN model reached 0.1297 and 0.0493, respectively after 6 epochs.

Figure 6: Comparative graph of SVM and CNN accuracy on test data. The SVM model achieved an accuracy of 80.5% whereas the accuracy of the CNN model was 98.75% with 0.0251 loss.

Results II

Figure 7: AUC-ROC curve for SVM Classifier with AUC score of 0.838 indicating that the model is relatively good at distinguishing between positive and negative class values.

Discussion

Takeaways

- This work proposes a novel convolutional neural network model for the classification of turnip leaves with early-stage Anthracnose infection.
- The high accuracy(98.75%) of the deep learning CNN model compared to that of the supervised learning SVM model(80.50%) demonstrate the efficacy of deep learning approaches for image-based plant disease diagnosis.

Future Directions

- CNN model can be developed into a real-time disease monitoring system to identify Anthracnose when it first appears in turnip leaves, enabling effective disease management of agricultural produce in farms worldwide.
- Additional steps to be taken include refining the model by implementing k-fold cross validation or experimenting with other activation functions such as SoftMax.
- Future studies should also be warranted to apply computer vision to other diseases that affect the turnip or to other plants of agricultural importance.

Acknowledgements

I thank Dr. Claudia Stone of California State University East Bay for her guidance in helping me with the research project design.