Plant Classification and Diseases Identification Using Image Segmentation

Mehrnoush Famil Dardashti [1] **Gabriel Torres** [1] **Sumeet Shanbhag** [1]

Abstract

This paper presents a plant image classification scheme that uses a combination of Unet-based image segmentation and a convolutional neural network (CNN) architecture for the actual classification. The first step of the proposed approach is to segment the plant leaves from the background using a modified Unet architecture, which is a popular deep-learning model for image segmentation. The segmented leaves are then preprocessed and fed into a CNN architecture for the actual classification. The CNN architecture consists of multiple convolutional layers, followed by pooling and fully connected layers, which enable the model to learn the complex features necessary for accurate classification. To evaluate the proposed approach, experiments were conducted on a publicly available Plant village dataset. The results show that the proposed approach achieves high accuracy in classifying different plant species.

1. Introduction

The percentage of crop yield lost to plant diseases each year varies depending on the crop, the region, and the specific plant disease. According to the Food and Agriculture Organization (FAO) of the United Nations, plant diseases and pests are responsible for the loss of up to 40% of global food crops each year. In some cases, the percentage of crop yield lost to plant diseases can be much higher. For example, some estimates suggest that up to 80% of banana crops worldwide are at risk from the fungal disease known as Panama disease, also called Fusarium wilt, which can cause complete crop loss. It's worth noting that losses due to plant diseases can also have significant economic impacts beyond just the loss of crops, including increased costs for control measures and reduced market access for farmers. Early detection of plant diseases remains challenging due to lacking lab infrastructure and expertise.

1.1. Research contributions

The main contributions of this work are given as follows:

- The development of an efficient plant image classification scheme that utilizes a U-Net-based image segmentation approach followed by a CNN architecture for classification.
- The use of the U-Net-based segmentation approach to accurately segment the plant images and generate binary masks for further processing.
- The demonstration of the effectiveness of the proposed approach through experimentation and comparison with existing methods.
- The potential application of the proposed approach in real-world scenarios, such as early detection and prevention of plant diseases for increased crop yields.

2. Related Work

Prior deep learning (DL) schemes were repopularized; the work of Malla et al. in (Mallah et al., 2013) presents a conventional machine learning approach that applies the KNN-classifier as a baseline algorithm for plant species identification. One-hundred plant species leaves data set developed in (Mallah et al., 2013) consists of a set of binary images (masks) of leaf samples without their corresponding color images as it can be seen in Fig.1 As part of the work done in (Mallah et al., 2013) three principal characteristics were extracted from each image: shape, margin, and texture. These features are given as 64-element vectors corresponding to each image. These vectors are presented as consecutive descriptors for shape and histograms for texture and margin. As a result, three distinct files are created, one for each feature problem (Mallah et al., 2013). There are 1600 samples with 16 samples per leaf class (100 classes) and no missing values.

In (Mallah et al., 2013), the shape, margin, and texture features are utilized for training some K-NN classifier variants. Moreover, different levels of model complexity are assessed by including or discarding features during the

[1]Worcester Polytechnic Institute. Correspondence to: Gabriel Torres <gatorresarosemen@wpi.edu>.

CS/DS 541 Deep Learning Final Project, Worcester Polytechnic Institute.

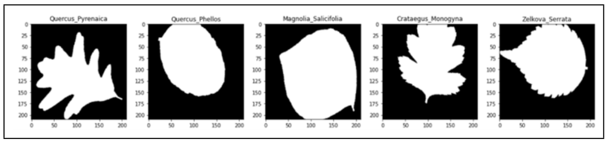

Figure 1. Binary image (masks) sample of plant specimen available in One-hundred plant species leaves data set

modeling phase. The performance of these algorithms is evaluated through the convergence of the mean logarithmic of the probability estimation (EEL), which resembles the cross-entropy loss function previously studied in this course. In addition, a more intuitive performance metric is given from the model's accuracy (ACC). Results reflect a very promising performance in the order of -0.956 and 96% for ELL and ACC, respectively. In (Gomaa, 2023), a CNN classification model is constructed for plant disease identification. In contrast with the work presented in (Gomaa, 2023), full-colored pictures of the 38 plant species for classification purposes. In Fig.(Gomaa, 2023), it can be observed that pictures of the same plant will show different patterns depending on the carried disease. In this case, a vectorial map of the pixels of each picture is used as a feature to train a convolutional neural network (CNN) utilizing TensorFlow and Keras Python libraries. Similarly, as done in (Mallah et al., 2013), logarithmic loss and model accuracy of 86% are used as metrics to evaluate the neural network's performance. The discrepancy in terms of accuracy with the results mentioned in (Mallah et al., 2013) and (Gomaa, 2023) indicate that significant information can be obtained from binary images, such as edges and corners and gradually build up to more complex features, such as shapes and objects. In addition, Mask images provide a localization component that detects relevant information within an image. For instance, during the last three years, binary images have been used to identify particular lesions on lung surrounding tissues caused by COVID-19. The work shown in the repository (Vitali, 2020) introduces a series of instructions for training a lung image segmentation model. The code in this repository is implemented using the PyTorch deep learning framework. The model architecture is based on a U-Net (Ronneberger et al., 2015), a convolutional neural network commonly used for segmentation tasks. The model is trained on a dataset of 2D scans with corresponding lung segmentation masks. The main objective of this project is to provide a plant image segmentation framework that allows the capture of relevant spatial characteristics of full-colored plant images for species classification and disease detection. For this purpose, a similar unit-based image segmentation framework as in (Vitali, 2020) is developed in order to concatenate binary images from the full-colored Plant Village dataset in (Gomaa, 2023) utilizing a DL scheme. Resultant Binary images will be considered inputs for a CNN-based image classification algorithm.

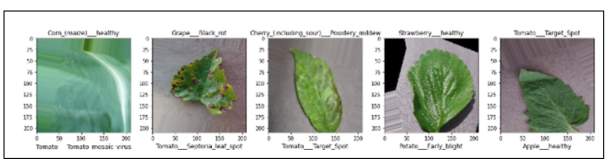

Figure 2. Sample of plant disease identification data set utilized in Gomaa et al.

3. Proposed Method

3.1. Data Description

This project uses the **Plant village dataset** found in (Gomaa, 2023).The Plant Village dataset is a collection of over 54,000 high-quality images of 14 different crop species, including tomato, potato, apple, and grape. Each image is associated with one of 38 different classes, representing various plant diseases or healthy conditions.

3.2. Computer Vision

The Python OpenCV library provides several functions for image thresholding, a process that converts grayscale or color images into binary images. In this process, each pixel in the image is compared to a threshold value and is assigned a binary value (0 or 255) based on whether it is above or below the threshold value. The OpenCV library is used in this project to convert the RGB images into segmented images with a black background, but it failed to give decent results for all the images, as shown in Fig.3. A qualitative study of the results of this computer vision scheme showed an efficient performance of this strategy on the tomato healthy leaves images. Consequently, the obtained tomato healthy binary images are used as the ground truth for training the UNET (Ronneberger et al., 2015) image segmentation strategy.

3.3. U-NET Image Segmentation

The U-net is an advanced deep learning architecture designed for image segmentation, focusing on biomedical image analysis (Ronneberger et al., 2015). Its name is derived from its U-shaped network architecture, distinguishing it from conventional CNN models. In contrast to standard CNN models, U-net employs convolutional layers to upsample or combine feature maps into a complete image. This project introduces the conventional U-Net architecture utilized in (Vitali, 2020) using the **tomato healthy** as a baseline training subset. In addition, to reduce the number of channels of each image, OpenCV is newly used to obtain grayscale images for faster training performance. The architecture utilized in this work follows the same structure

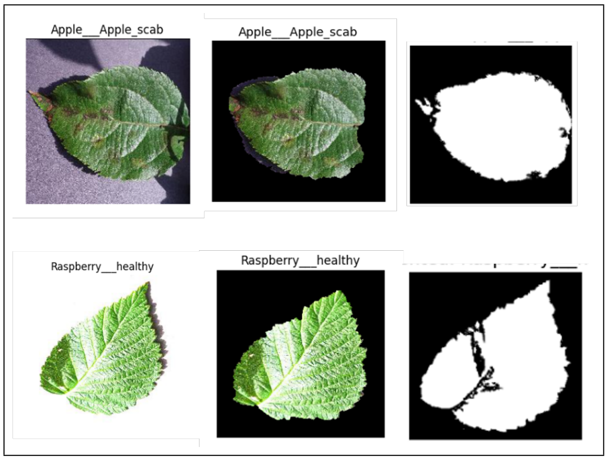

Figure 3. OpenCV image processing results.

used in (Vitali, 2020) and (Chowdhury et al., 2021) as it can be observed in Fig.4.

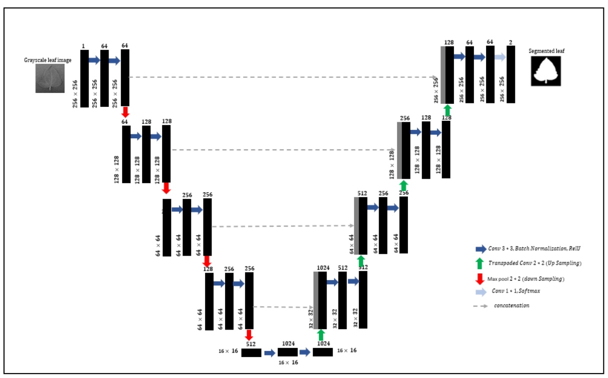

Figure 4. Architecture of original U-Net.

3.4. CNN Image Classification

In this section, a conventional CNN-based classification scheme is presented. The model obtained in the previous U-NET segmentation framework is used to predict binary images from full-colored images on the rest of the dataset. However, since different disease information can not be characterized using binary inputs, the color images are superimposed with their respective binary image to remove additional noise from the background and other nonrelevant features in each image input. This process is expected to enhance the training process regarding computational cost and information capturing, as illustrated in Fig.5.

3.5. Regularization

Since the dataset is not balanced and does not have a similar number of images for the different categories, training with an imbalanced dataset can produce a biased model. Thus, data augmentation can help to provide a similar number of images in the various classes, which can provide reliable results, as stated in many recent publications. Three augmentation strategies (rotation, scaling, and translation) were utilized to balance the training images.

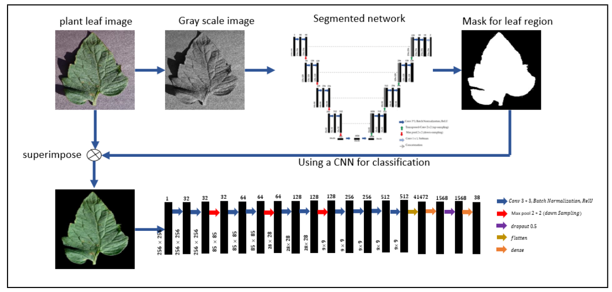

Figure 5. Architecture of CNN classification structure.

4. Experiment

- Image Segmentation Experiments: The performance of the U-NET image segmentation algorithm is tested with respect to both validation and testing data regarding binary cross entropy loss function and accuracy. Furthermore, a Mobilenetv2 pre-trained neural network (Sandler et al., 2018) is introduced as the backbone of the current U-Net scheme to strengthen the robustness of the structure and facilitate the training process. Results of both schemes are compared with respect to performance metrics binary cross entropy and classification accuracy.

- Image Classification Experiments: The performance of the CNN classification architecture is tested on the original full-colored dataset, the predicted segmented image dataset, and the full-colored augmented dataset in terms of the sparse categorical cross entropy loss function and model accuracy.

5. Results

5.1. Image Segmentation Results

In Fig.6 are illustrated the performance metrics at each training epoch of the conventional U-Net image segmentation. Whereas the Mobilenetv2 U-Net segmentation outcomes are depicted in Fig.7. Results of both training schemes are summurized in Table.1 where it can be noticed that as expected the inclusion of the pre-trained Mobilenetv2 backbone outperforms the conventional U-Net model.

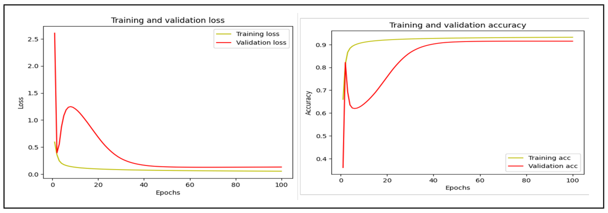

Figure 6. Conventional U-Net Image Segmentation Binary Cross Entropy and Accuracy Time Evolution.

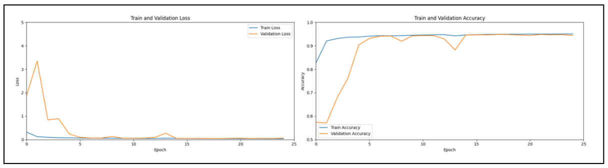

Figure 7. U-Net with Mobilenetv2 Backbone Image Segmentation Binary Cross Entropy and Accuracy Time Evolution.

Model	Train ACC	Validation ACC
U-NET	0.9321	0.9148
U-NET-Mobilenetv2	0.9506	0.9439

Table 1. Image Segmentation Results.

5.2. CNN-Classification Results

The classification performance metrics for the segmented images are given in Fig.8. Classification results utilizing full-colored images and augmented-colored images are observed in Fig9 and Fig.10. Accuracy metrics are presented in Table.2 show that the CNN algorithm has a promising performance in terms of classification accuracy for the colored augmented data input.

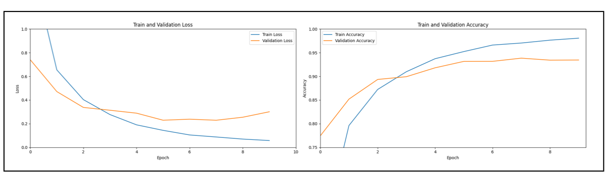

Figure 8. Classification with segmented photos.

Model	Train ACC	Validation ACC
Segmented	0.9807	0.9345
Full-Colored	0.9792	0.9489
Color-Augmented	0.9831	0.9681

Table 2. CNN Classification Results.

6. Discussion

The main challenge found regards to computational power available which limited the training process of the U-Net-based segmentation scheme to a significantly small subset

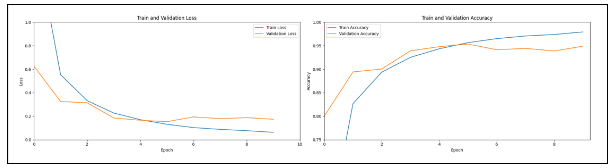

Figure 9. Classification with colorful photos

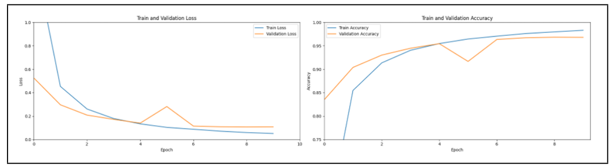

Figure 10. Classification with augmented colorful photos.

of data samples. This prevents the model from being able to generalize different crops with different leaf patterns. In addition, it must be noticed that image inputs can be easily segmented since images only contain photos of the leaves of each crop, it may be interesting to study how the performance of the algorithm is impacted if the input image contains additional background noise or other types of content.

7. Conclusions and Future Work

U-Net-based segmentation strategy has shown to have a decent effectivity in characterizing a picture's contour and in some sense obtained similar characteristics as the ones found in (Mallah et al., 2013) work. In fact, another interesting approach that can be performed is to utilize segmented images from UNET as inputs for the ML scheme developed in (Mallah et al., 2013) for plant classification. The developed CNN-based architecture for classifications has shown to be more efficient when full-color images than with segmented pictures. However, notice that the plant village dataset only contains images with only specific leaves pictures. Thus, the performance of this scheme may vary if the input image contains additional background noise or some other kind of content. U-Net-based segmentation provides an interesting baseline in order to characterize only relevant information for plant and disease identification purposes. Further work must be done in terms of data processing. For example, to consider a larger and more varied dataset for image segmentation UNET training and to utilize a more refined DL scheme, such as MRCNN, that provides a more efficient information detection performance.

References

Chowdhury, Muhammad EH, Rahman, Tawsifur, Khandakar, Amith, Ayari, Mohamed Arselene, Khan, Aftab Ullah, Khan, Muhammad Salman, Al-Emadi, Nasser, Reaz, Mamun Bin Ibne, Islam, Mohammad Tariqul, and Ali, Sawal Hamid Md. Automatic and reliable leaf disease detection using deep learning techniques. *AgriEngineering*, 3(2):294–312, 2021.

Gomaa, Ahmed. Plant diseases classification cnn, kaggle repository, version 7. 2023. URL https://www.kaggle.com/code/ahmedgomaa22/plant-diseases-classification-cnn-model.

Mallah, Charles, Cope, James, Orwell, James, et al. Plant leaf classification using probabilistic integration of shape, texture and margin features. *Signal Processing, Pattern Recognition and Applications*, 5(1):45–54, 2013.

Ronneberger, Olaf, Fischer, Philipp, and Brox, Thomas. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241. Springer, 2015.

Sandler, Mark, Howard, Andrew, Zhu, Menglong, Zhmoginov, Andrey, and Chen, Liang-Chieh. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018.

Vitali, Petsiuk. Lung-segmentation-2d. 2020. URL https://github.com/imlab-uiip/lung-segmentation-2d#readme.