Leaf Disease Segmentation

A micro deep learning project report submitted in partial fulfilment of the requirement for the award of the degree of

Bachelor of Engineering

in

Electronics and Computer Engineering

Submitted By

Fateh Singh (102215013)

Rushil Chatley (102215261)

Vaibhav Gupta (102215113)

Team No: 33

Under Supervision of

Dr. Gaganpreet Kaur (Assistant Professor)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Department of Electronics and Communication Engineering

THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY, PATIALA, PUNJAB

December 2025

Leaf disease segmentation

Abstract

Monitoring plant status is one of the most important tasks in digital agro. It is important to determine the type of disease, as well as its location, that can be used for calculation of percentage of plant damage. Our task is to fulfil these two tasks through classification, segmentation and class activation map. We also present the updated Plant Pathology 2020 dataset (4) with ground truth masks of leaf diseases.

1. Introduction

Agricultural productivity is something on which economy highly depends. This is the one of the reasons that disease detection in plants plays an important role in agriculture field, as having disease in plants are quite natural. The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops.

Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance. Whereas if automatic detection technique is used it will take less efforts, less time and become more accurate.

In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases. Image processing is used for measuring affected area of disease and to determine the difference in the colour of the affected area.

Plant monitoring may include classification to determine the type of disease, as well as determining the localization of the disease and the percentage of plant damage using segmentation. The latter is much more difficult to do, as there are simply no labelled datasets in the industry. In our work, we present an updated Plant Pathology 2020 dataset in which we managed to get ground truth masks of leaf diseases. We also offer a solution to the segmentation problem, as well as an alternative method for determining the localization of the disease without mask annotated dataset using classification and class activation map.

2. Related work

In one of the related works (5) considering method of recognizing apple leaf diseases through region-of interest-aware deep convolutional neural network is proposed in this paper. The primary idea is that leaf disease symptoms appear in the leaf area whereas the background region contains no useful information regarding leaf diseases. To realize this idea, two subnetworks are first designed. One is for the division of the input image into three areas: background, leaf area, and spot area indicating the leaf diseases, which is the region of interest (ROI), and the other is for the classification of leaf diseases. The two subnetworks exhibit the architecture types of an encoder–decoder network and VGG network, respectively; subsequently, they are trained separately through transfer learning with a new training set containing class information, according to the types of leaf diseases and the ground truth images where the background, leaf area, and spot area are separated.

Next, to connect these subnetworks and subsequently train the connected whole network in an end-to-end manner, the predicted ROI feature map is stacked on the top of the input image through a fusion layer, and subsequently fed into the subnetwork used for the leaf disease identification. The experimental results indicate that correct recognition accuracy can be increased using the predicted ROI feature map. It is also shown that the proposed method obtains better performance than the conventional state-of-the-art methods: transfer-learning-based methods, bilinear model, and multiscale-based deep feature extraction and pooling approach.

In this experiment, the mean accuracy of the ROI subnetwork, defined as the ratio of correctly classified pixels to total pixels for each class, was 86% and mean Iowa (Intersection over Union), also known as the Jacquard similarity coefficient, is 69%. The VGG subnetwork, which making classification, shows the experimental result of correct recognition accuracy of 74.7 %.

Instance Segmentation of Biological Images Using Harmonic Embedding's (2) approach describes each object instance using an expectation of a limited number of sine waves with frequencies and phases adjusted to particular object sizes and densities. At train time, a fully-convolutional network is learned to predict the object embedding's at each pixel using a simple pixel wise regression loss, while at test time the instances are recovered using clustering in the embedding space.

3. Data and methods

We have captured high-quality, real-life RGB images of apple leaves with diseases from Plant Pathology 2020 (4). Originally dataset contain 3642 images with class labels, there are 4 classes namely healthy, rust disease, scab disease, both diseases. We managed to get ground truth mask disease annotations for 1291 images from dataset (400 images for each of healthy, rust, scab and 91 images for both diseases class). We used only this 1291 images in our work.

The complexities of the dataset were increased by including:

- Imbalanced dataset of different disease categories
- Non-homogeneous background of images
- Images taken at different times of day
- Images from different physiological age of the plants
- Multiple diseases in the same image
- Different focus of the images

Examples of pictures are follows: healthy leaves (Figure 1), apple scab (Figure 2), cedar apple rust (Figure 3), and both diseases (Figure 4). The dataset was randomly split into training and stratified test set of 80 % and 20%, respectively, such that both datasets have all four disease categories.

Figure 1. Health

Figure 2. Health

Figure 3. Rust

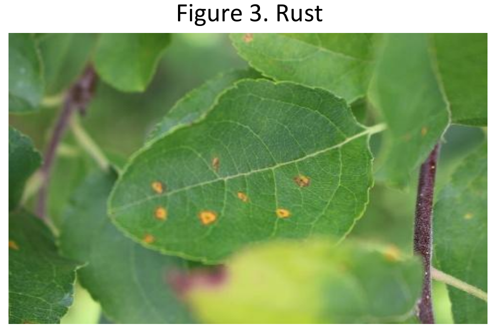


Figure 4. Both Disease

3.1. Disease classification using a standard convolutional neural network

All leaf disease images were adjusted so that the length and width of the image were the same, which were resized to 352×512, because it was hard to use full size images due to speed and memory restrictions. For data augmentation we used horizontal and vertical flips and random rotation.

So we have trained a convolutional neural network on this dataset for classification of apple scab, cedar apple rust, leaves with both diseases, and healthy leaves. Specifically, we took a ResNet18 (1) network and fine-tuned the network weights on our annotated disease dataset. We use Cross entropy loss with weights to reduce influence of imbalances of data on the results.

We also decided to do classification with mask, when the input of net will be sent not only the image, but also the mask as additional channel and classification with only mask. Motivation for it is to understand quality of annotated ground truth mask. In case image + mask we also used pre-trained ResNet18, but change not only last but first layer too (because now we have 4 input channels). The we freeze all layer besides new layers. Firstly, we train network with freeze layers, then unfreeze them and train whole net. In case when input is only mask we used ResNet18 without any pre-trained weights.

3.2. Class activation map

Class activation map method allows to interpret predictions of convolutional neural networks (6). A typical convolutional network that is used for classification consists of a fully convolutional part followed by an average pooling and a fully connected layer:

$out_{con} = Cong(input) \in R_{c \times h \times w}$ $out^{pool} = AveragePooling(out^{conv}) \in R^{c}$ $out_{logit} = W_{T} \times out_{pool} \in R_{m}, W \in R_{c \times m}$

Here we can consider j-th column of a weight matrix $W \in R^{c \times m}$, where m is a number of classes and c is a number of channels in out^{conv} tensor, as importances of each of c channels for j-th class. Using this assumption we can consider a weighted sum of channels of out^{conv} as a map of importances for j-th class.

$$CAM^{j} = \sum_{i=1}^{c} |w_{ij}||out_{i}^{conv}|$$

Or for each pixel:

$$CAM_{kl}^{j} = \sum_{i=1}^{c} |w_{i}||out_{ikl}^{conv}|$$

Here we need just to scale CAM^{j} on an interval $(0,1)$ and resize to size of the initial image. Class activation map helps to explore behavior of a convolutional neural network model, for example sometimes it can help to find a reason of a wrong classification result. Also it can be used as a segmentation model that helps to define locations of class characteristic features of a class. This is very useful for our problem, because usually there is no data for training segmentation models.

3.3. Segmentation

For segmentation we use Unet (3) architecture with number base channels equal to 64. We also resize images and use the same augmentation as for classification. We experimented with different learning rates, optimizers, weights for loss and schedulers. We also trained net with same parameters for 5, 10, 30, 50, 100 percent of train dataset and same validation set to find out dependence of size of dataset on performance.

4. Results

4.1. Classification

The best result for classification are follows:

Table 1. Classification result

Data	Accuracy on val. set	Loss on val. set
Images	0.946	0.232
Images+gt masks	0.965	0.152
gt masks	0.954	0.212

From results we can conclude that ground truth mask annotation for our dataset is good, but far from perfect, as evidenced by manual viewing of the annotation. In most cases we got bad ground truth mask for scab disease because it was much more complex to annotate this disease and we asked annotators to skip those leaves where they are not sure is it disease or not. You can see some examples of bad grout truth mask of Figures 5, 6. On Figure 5 you can see that annotation is quite rude and inaccurate. On Figure 6 there are a lot of not annotated spots. But such quality of ground truth mask leads to interesting results that we can observe further is segmentation result part.

4.2. Class activation map

Trained ResNet-18 classification model (when input was only image) from previous section was used to obtain class activation maps using the algorithm described above. The model accepts a colored image as an input and returns scores for each class. The model shows good results on leaves with diseases, examples are shown on Figures 9, 10.

Figure 5. Bad ground truth mask example.

Figure 6. Bad ground truth mask example.

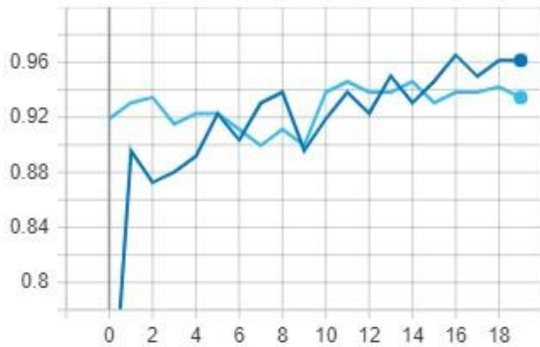


Figure 7. Classification accuracy curve on val set. *Light blue images*, *Dark blue - images+gt mask*

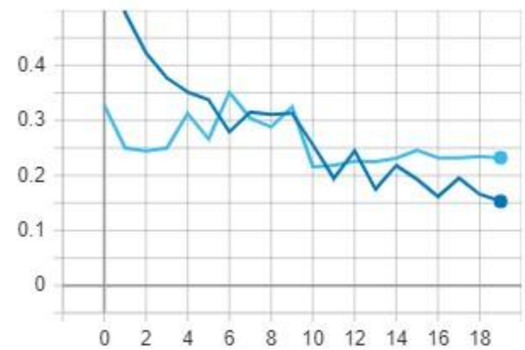


Figure 8. Classification loss curve on val set. *Light blue - images*, *Dark blue - images+gt mask*

There is a problem with healthy leaves. They are classified correctly, but because calculated CAM values are not absolute and scaled on an interval $(0,1)$ even healthy leaves will have activations on some regions without diseases. Example

Figure 9. Class activation map, Scab

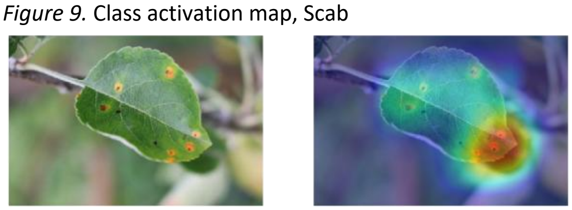


Figure 10. Class activation map, Rust

is shown on 11.

Figure 11. Class activation map, Healthy

4.3. Segmentation

You can see segmentation results for rust leaves disease (Figure 12), scab leaves disease (Figure 13), healthy leaves (Figure 14), both diseases (Figure 15). As you can see result for healthy and rust around perfect, predicted mask completely or almost completely match with ground truth. More interesting result you can see on scab and both diseases example. We found that when using a 100 percent of train dataset and selecting the right hyperparameters, the model often correctly segments those areas that were not marked as ground truth. In most cases, this is true for scab disease because the fact we mentioned above - we have much worse ground truth annotation for this class, because it was much more complex disease for visual perception and we asked annotators to skip those leaves where they are not sure is it disease or not. On the Figure 13 you can see that our model correctly segment disease in the up right corner. For both diseases example (Figure 15) you can see correct segmentation of unlabeled leaf at the down left corner.

As for errors and incorrect segmentation, you can see several examples in figures 16 and 17. In first case, we see false positive mask for branches in the second case see false positive for bright apple.

Figure 12. Rust. *left - image*, *middle - mask*, *right - predicted mask*

Figure 13. Scab. *left - image*, *middle - mask*, *right - predicted mask*

Figure 15. Both diseases. *left - image*, *middle - mask*, *right - predicted mask*

Figure 16. Bad prediction result. *left - image*, *middle - mask*, *right - predicted mask*

On the Figure 18 and 19 you can see IoU and true positive rate dependences on train dataset size respectively. This curves match with fact that we mentiones above about frequent correct segmentation for not annotated regions. IoU

Figure 17. Bad prediction result. *left - image*, *middle - mask*, *right - predicted mask*

increase for rust always when train size increase, this is not true for scab for 100 percent train dataset size, but in order to make sure that out model work better for ground truth annotation on entire dataset we calculate true positive rate and as we can see - full size train dataset give here the best result. In addition from curves we can see that with only 30% of train dataset we already

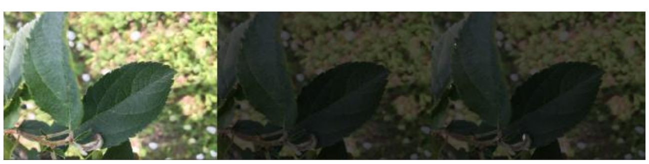
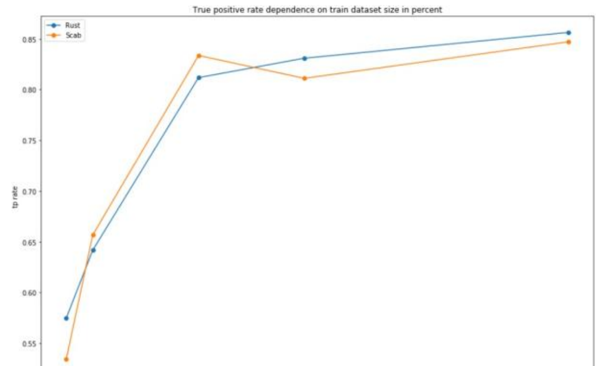


Figure 14. Healthy. *left - image, middle - mask, right - predicted mask*

can achieve quite good results.

Figure 18. Intersection over union dependence on train dataset percent. *Blue - Rust*, *Orange - Scab*

Figure 19. True positive rate dependence on train dataset



percent. *Blue - Rust*, *Orange - Scab*

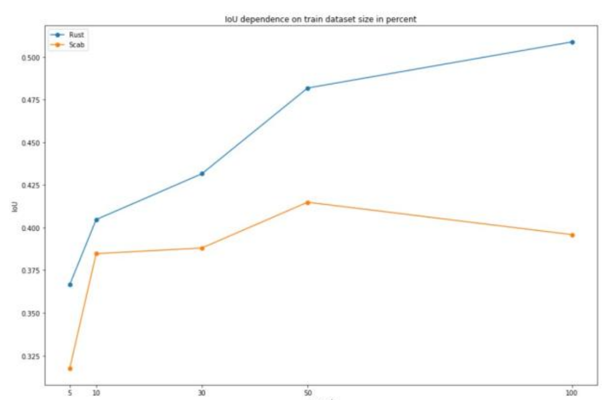
4.4. Comparison with state-of-art

To compare our results with state of the art we took best classification result in Kaggle competition, because there is no article with classification result with this dataset (because this is a new dataset) and since in the kaggle competition the entire dataset was used we trained our network on whole dataset.

Table 2. Comparison of result

Metric	Our approach	Best kaggle result
Accuracy of classification	0.968	0.984

For segmentation we comparing our approach with method from article (5). But it should be noted that in this work one of the segmented classes was the whole



sheet, which, of course, is an easier task, because of this we have less value of IoU. We segmented only disease.

Table 3. Comparison of result

Metric	Our approach	ROI approach

Mean IoU	0.45	0.69

5. Conclusion

We obtain good and interesting results for classification, class activation map and segmentation.

For classification we achieve quite good result based on the traditional ResNet18 model, that then was used for class activation map. We also check quality of dataset by training the network with input as images with ground truth mask, which improved the classification accuracy of the model on the plant leaf disease dataset.

Class activation maps showed very interesting results. We are able to locate diseases on photos without even using segmentation models. This can be very helpful if there is no segmentation data. Also in this case we need much less computing power and can use the model on embedded devices in real time.

Segmentation gives us good and interesting results, it can be successfully used to find biggest disease damage locations of plant, as well as to determine the level of this damage. There is a huge field for working for the future, both with the annotations we presented for the part of Plant Pathology 2020 dataset and with the model that can be improved and also can be used for segmentation of other diseases using fine-tuning.

References

[1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.

[2] Victor Kulikov and Victor S. Lempitsky. Instance segmentation of biological images using harmonic embeddings. *CoRR*, abs/1904.05257, 2019.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[4] Ranjita Thapa, Noah Snavely, Serge Belongie, and Awais Khan. The plant pathology 2020 challenge dataset to classify foliar disease of apples, 2020.

[5] Hee-Jin Yu and Chang-Hwan Son. Apple leaf disease identification through region-of-interest-aware deep convolutional neural network. *CoRR*, abs/1903.10356, 2019.

[6] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude` Oliva, and Antonio Torralba. Learning deep features for discriminative localization. *CoRR*, abs/1512.04150, 2015.