SEMANTIC SEGMENTATION FOR PLANT DISEASE USING DEEP LEARNING

SHU SHI HAO

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic Systems)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

JANUARY 2020

DEDICATION

This thesis is dedicated to my parents, who taught me never to give up and always strive to be the best. It is also dedicated to my siblings, who provided me with moral support throughout the entire length of the project.

ACKNOWLEDGEMENT

I would like to take this opportunity to thank all people who have helped me throughout the entire duration of this project. First of all, I would like to express my sincere gratitude to my project supervisor, Dr Musa bin Mohd Mokji and a PhD student, Aliyu Muhammad Abdu for their guidance, wisdom and brilliant ideas which they shared during the course of this project. I would also like to thank them for spending their valuable time coming all the way out to Penang to have discussion with me regarding the project and provided me with invaluable advice. Their guidance and insight have been the main reason I am able to complete my project.

Secondly, I would like to thank my panel of examiners Assoc Prof Muhammad Mun'im Ahmad Zabidi and Dr Shahidatul Sadiah Abdul Manan for providing constructive feedback during the viva session.

Last but no least, I would like to thank my parents, family members and friends for their encouragement and moral support thorughout the duration of this project.

ABSTRACT

In the era of artificial intelligence, various applications are using machine learning to solve some of the engineering problems and ease the people work. Plant disease is important to be detect as early as possible because the early we found out there is a disease, the lesser in loss of crops from plant disease which may impact to the economic and increase of price for crops. Numerous researches have been reported in the literature that employed different aspects of machine learning methods for plant disease detection, with the majority of them are focusing on the plant leaves. There are two common and destructive foliar diseases that are early blight and late blight. Early blight infection usually associated with plant physiological maturity and fruit load and late blight can infect and devastate the plants at any developmental stages. Since the leaves are found to be the most commonly observed part for detecting an infection. Segmentation technique is a basic and easy way to classify and estimate the severity of the diseases because it works well with plant disease detection since the infected leaf area shows significant color differences from its original color. Feature extraction is an essential step before the segmentation process that determines the applicability of every machine learning model. Convolution Neural Network (CNN) is a method that is gaining popularity to solve the feature extraction problem since it can automatically extract the features directly from the input images. Hence, this project aims to utilize the semantic segmentation with CNN through transfer learning from the VGG16 network to segment the plant leaf into healthy, necrotic and symptomatic regions. There are 1200 samples to be labeled as healthy, necrotic and symptomatic and classified into three severity levels to serve as the training material in the supervise training of CNN. To observe the performance of the training, normal training and optimized training will be compared in terms of accuracy and efficiency. Lastly, a deep learning model will be developed and is capable of recognizing and labeling the given leaves regions whether it is healthy, necrotic and symptomatic. The deep learning model able to achieve the global accuracy of 93% and IoU of 67.83% with 1200 data samples which segmented into 4 different classes.

ABSTRAK

Dalam era kepintaran buatan, pelbagai aplikasi menggunakan pembelajaran mesin untuk menyelesasikan masalah-masalah kejuruteraaan and memudahkan orang. Penyakit tumbuh-tumbuhan penting untuk dikesan secepat mungkin kerana kita boleh mengurangkan kesan kehilangan tanaman dari penyakit tumbuh-tumbuhan and kenaikan harga tanaman kalau penyakit didapati pada awalnya. Banyak penyelidikan telah dilaporkan dalam kesusateraan yang menggunakan pelbagai aspek kaedah pembelajaran mesin untuk mengesanan penyakit tumbuh-tumbuhan. Kebanyakan mereka memberi tumpuan kepada daun tumbuh-tumbuhan. Terdapat dua penyakit foliar yang biasa dan merosakkan tumbuh-tumbuhan. "Early Blight" yang buasanya dikaitkan dengan kematangan fisiologi tumbuh-tumbuhan dan hasil buah. "Late Blight" dapat menginfeksi dan menghancurkan tumbuh-tumbuhan pada mana-mana peringkat perkembangan. Oleh kerana daun didapti adalah bahagian yang paling biasa diperhatikan untuk mengesan jangkitan. Teknik segmentasi adalah satu cara asas dan mudah untuk mengklasifikasikan dan menganggarkan keterukan penyakit kerana ia dapat mengesan penyakit tumbuh-tumbuhan. Kawasan daun yang dijangkiti menunjukkan perbezaan warna yang signifikan dari warna asalnya. Pengekstrakan ciri adalah satu langkah yang perlu sebelum proses segmentasi yang menentukan pemakaian pembelajaran mesin. "Convolutional Neural Network (CNN)" adalah satu kaedah yang semakin popular digunakan untuk menyelesaikan masalah pengekstrakan ciri kerana ia boleh mengeluarkan ciri-ciri secara langsung dari gambar yang dibagikan.Oleh itu, projek ini bertujuan untuk menggunakan segmentasi semantik dengan menggunakan CNN melalui pembeljaran pemindahan dari rangkaian VGG16 untuk membahagikan daun tumbuh-tumbuhan kepada kawasan sihat, nekrotik dan symptomatik. Terdapat 1200 sampel yang dilabelkan sebagai kawasan sihat, nekrotik dan symptomatik dan diklasifikasikan ke dalam tiga tahap keparahan untuk menjadi bahan latihan dalam latihan penyeliaan CNN. Untuk melihat prestasi latihan, latihan biasa dan latihan yang optimum akan dibandingkan dari segi ketepatan dan kecekapan. Akhir sekali, satu model "Deep Learning" dibuat dan ia dapat mengenali, mengabelkan kawasan daun yang diberikan sama ada sihat, nekrotik dan symptomatik. Model pembelajaran mendalam dapat mencapai ketepatan global sebanyak 93% dan IoU daripada 67.83% dengan 1200 sampel data yang dibahagikan kepada 4 kelas yang berlainan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CNN	-	Convolutional Neural Network
R-CNN	-	Region-based Convolutional Neural Network
ReLU	-	Rectified Linear Units
SVM	-	Support Vector Machine
RGB	-	Red Green Blue
ANN	-	Artificial Neural Network
IoU	-	Intersection-over-Union
BF	-	Boundary F1

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Problem Background

Plant disease usually defined as abnormal growth of a plant. Disease is the result of some disturbance in the plant normal life process. It may cause by living organisms or non-living environmental conditions such as weathers, soil compaction, and so on [1]

Plant disease is a major threat to food production and plant diversity, but rapid identification is still difficult in many places in the world due to the lack of required infrastructure. In this smartphone era, the combination of rapid increment of global smartphone usage and the advanced technology in computer vision made smartphone-assisted disease diagnosis by deep learning is possible [2]. Modern technologies have helped human society to produce enough food demand for more than 7 billion people. But, food security is still threatened by some factors including climate change [3], the decline in pollinators [4], plant diseases [5], and others. Therefore, plant diseases require monitoring mechanisms to take appropriate measures to ensure the food production [6]. Imagine if no appropriate action taken on the plant diseases, this may become a disaster for the 7.7 billion of the world's human population. Plant diseases are not only a threat to food security globally, but they can also bring into disastrous consequences for smallholder farmers whose livelihoods depend on the healthiness of their crops. In the developing country, more than 80 percent of the agricultural production is generated by smallholder farmers [7], and yield loss from the reports is more than 50 percent due to pests and diseases commonly [8]. Furthermore, the high density of hungry people lives in smallholder farming households [9], making smallholder farmers a group that is vulnerable to pathogen-derived disruptions in supply of food particularly.

Different efforts have been developed to reduce and prevent loss of crops due to the attack of diseases. In the past decade, historical approaches of the widespread application of pesticides have increasingly been supplemented by integrated pest management (IPM) approaches [10]. Identifying a disease correctly is a very crucial step for efficient disease management when it first appears. In the history, identification of disease has been supported by agricultural extension organizations or even other institutions such as local plant clinics. Recently, the efforts have been supported additionally by providing information for disease diagnosis online, leveraging the improvement of high-speed Internet services worldwide. With the mobile phones have proliferated in recent years, taking advantage of the historically unparalleled rapid uptake of mobile phone technology all around the world [11].

Smartphones are a new rising approaches to help in identifying the diseases due to their computation ability is rising, high resolution display and wide range of built-in sets of accessories such as advanced cameras and etc. From the estimation, there will be 5 to 6 billion of smartphones in the global by 2020. From the previous survey by the end of 2015, there are 69 percent of the world's population had access to mobile broadband coverage and 47 percent penetration, a 12-fold increase since 2007 [11]. With the combined factors of the widely used smartphone, advanced cameras and high performance processors in the smartphone lead to a situation where disease diagnosis based on automated image recognition. From the theory, it is making sense that we can expect that technically feasible too. Previous work in [12] shows an example that demonstrate the technical feasibility by using a deep learning approach utilizing 54,306 images of 14 different crop species with 26 diseases including healthy made available through the project of PlantVillage [12].

1.2 Problem Statement

Plant disease should be detected as early as possible to help reducing the loss of crops from plant disease and prevent the increment of the crops price indirectly. From [13], plant leaves are the commonly observed part to detect an infection. Most of the existing work is identifying the disease based on the whole leaf image. It is done by extracting features like color and, texture, followed by identifying the disease based on the features.

In general, there are two common and destructive foliar diseases which are early blight and late blight. Early blight infection usually associated with plant physiological maturity and fruit load and late blight can infect and devastate the plants at any developmental stages. To differentiate these two types of disease, it is very hard for the existing work to identify the input as early blight or late blight since both of them have a very similar symptoms. Therefore, pathologist tends to identify, classify, quantify, and predict the plant disease based on the disease phenotype morphology. This is where the leaf is divided into the necrotic, symptomatic, and healthy areas. Hence, this project will adopt the existing pathologist approach which is to first segment the leaf into several meaningful areas.

Problem: Existing works are unable to identify early blight and late blight since both of them have very similar symptoms.

1.3 Objectives

The objective of this project is to utilize semantic segmentation approach to produce a more informative feature as identification of diseases. Semantic Segmentation describes the process of associating each pixel of an image with a class label. The plant leaf image will be segmented into three disease morphology areas, which are healthy, symptomatic, and necrotic, by applying deep Convolutional Neural Network (CNN) through transfer learning with VGG16 as a platform for informative features generated by semantic segmentation.

1.4 Scope of Study

The scope of this project is aimed at determining plant diseases from the images prepared. The focus will be given on 1200 samples of plant leaf images that are randomly selected from the tomato and potato leaves folder through the online platform, Plant Villages with more than 50,000 images. The 1200 image samples that are used to label and they will be used as the input of the transfer learning. There are three different disease morphology areas to be labeled necrotic, symptomatic, and healthy areas. The necrotic area indicates a plant cell with dead tissue or generated tissue where the area is darken and wilt. Besides that, symptomatic indicates some symptom on the plant leaves image. With the transfer learning method, VGG16 is being considered as pre-trained network to be used.

1.5 Thesis Outline

This project report consisted of five chapters. In Chapter 1, problem background, problem statement, objectives, and scope of this study are discussed. In Chapter 2, the background of deep learning and review of some related works are explained in detail. Chapter 3 presents the project methodology. The ways to conduct the research, the proposed algorithm and evaluation metrics are discussed in this chapter. After that, the project results are presented in Chapter 4. Detailed analysis for the results, discussions on the results obtained as well as comparison with existing works are presented in this chapter. Finally, the conclusion and recommendations for future work are discussed in Chapter 5.

REFERENCES

1. Coley-Smith, J. Plant pathology. *Physiological Plant Pathology*, 2008. 1(1): 83. ISSN 00484059. doi:10.1016/0048-4059(71)90043-9.
2. Mohanty, S. P., Hughes, D. P. and Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*, 2016. 7(September): 1–10. doi: 10.3389/fpls.2016.01419.
3. Kang, Y., Khan, S. and Ma, X. Climate change impacts on crop yield, crop water productivity and food security - A review. *Progress in Natural Science*, 2009. 19(12): 1665–1674. ISSN 10020071. doi:10.1016/j.pnsc.2009.08.001. URL `http://dx.doi.org/10.1016/j.pnsc.2009.08.001`.
4. Kuala Lumpur, P. o. t. I. S.-P. P. o. B. and Session, E. S. F. *Report of the Plenary of the Intergovernmental Science-PolicyPlatform on Biodiversity and Ecosystem Services on the work of its fourth session*. Technical report. 2016.
5. Strange, R. N. and Scott, P. R. Plant Disease: A Threat to Global Food Security. *Annual Review of Phytopathology*, 2005. 43(1): 83–116. ISSN 0066-4286. doi:10.1146/annurev.phyto.43.113004.133839. URL `https://doi.org/10.1146/annurev.phyto.43.113004.133839`.
6. Qi, J., Chen, J., Lafortezza, R. and Lin, Z. Remote Sensing for Ecosystem Sustainability. *Comprehensive Remote Sensing*, 2017: 186–201. doi:10.1016/b978-0-12-409548-9.10428-2.
7. Smallholders, F. S. and the Environment. UNEP. *Rome : International Fund for Agricultural Development (IFAD)*, 2013. URL `https://www.ifad.org/documents/10180/666cac2414b643c2876d9c2d1f01d5dd`.
8. Harvey, C. A., Harvey, C. A., Rakotobe, Z. L., Rao, N. S., Dave, R., Razafimahatratra, H., Rabarijohn, H., Rajaofara, H. and Mackinnon, J. L. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar Receive free email alerts when new articles cite this a. 2014. (October 2015). doi:10.1098/rstb.2013.0089.
9. Sanchez, P. A. and Swaminathan, M. S. Cutting world hunger in half. *Science*, 2005. 307(5708): 357–359. ISSN 00368075. doi:10.1126/science.1109057.
10. Ehler, L. E. Integrated pest management (IPM): definition, historical development and implementation, and the other IPM. *Pest Management Science*, 2006. 62(9): 787–789. ISSN 1526-498X. doi:10.1002/ps.1247. URL `https://doi.org/10.1002/ps.1247`.

11. Sanou, B. MDGs 2000-2015: ICT revolution and remaining gaps. 2015. URL https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf.

12. Hughes, D. P. and Salathe, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. 2015. URL http://arxiv.org/abs/1511.08060.

13. Kaur, S., Pandey, S. and Goel, S. Plants Disease Identification and Classification Through Leaf Images: A Survey. *Archives of Computational Methods in Engineering*, 2019. 26(2): 507–530. ISSN 18861784. doi:10.1007/s11831-018-9255-6. URL https://doi.org/10.1007/s11831-018-9255-6.

14. Rusk, N. Deep learning. *Nature Methods*, 2015. 13(1): 35. ISSN 15487105. doi:10.1038/nmeth.3707.

15. Bhatt, S., Patwa, F. and Sandhu, R. An Access Control Framework for Cloud-Enabled Wearable Internet of Things. *Proceedings - 2017 IEEE 3rd International Conference on Collaboration and Internet Computing, CIC 2017*, 2017. 2017-Janua: 328–338. doi:10.1109/CIC.2017.00050.

16. Bordes, A., Chopra, S. and Weston, J. Question Answering with Subgraph Embeddings. 2014. URL http://arxiv.org/abs/1406.3676.

17. Jean, S., Cho, K., Memisevic, R. and Bengio, Y. On Using Very Large Target Vocabulary for Neural Machine Translation. 2014. 000. URL http://arxiv.org/abs/1412.2007.

18. Sutskever, I., Vinyals, O. and Le, Q. V. Sequence to Sequence Learning with Neural Networks. 2014: 1–9. URL http://arxiv.org/abs/1409.3215.

19. Marr, B. Deep Learning Vs Neural Networks - What ' s The Difference ? 2019: 1–7. URL https://bernardmarr.com/default.asp?contentID=1789.

20. Strigl, D., Kofler, K. and Podlipnig, S. Performance and scalability of GPU-based convolutional neural networks. *Proceedings of the 18th Euromicro Conference on Parallel, Distributed and Network-Based Processing, PDP 2010*, 2010: 317–324. doi:10.1109/PDP.2010.43.

21. Radzi, S. A. and Khalil-Hani, M. Character recognition of license plate number using convolutional neural network. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011. 7066 LNCS(PART 1): 45–55. ISSN 03029743. doi:10.1007/978-3-642-25191-7_6.

22. Deshpande, A. A Beginner's Guide To Understanding Convolutional Neural Networks Part 2. 2016. i: 1–17.

23. Lecun, Y., Bottou, L., Bengio, Y. and Ha, P. Gradient-Based Learning Applied to Document Recognition In Intelligent Signal Processing. 1998. (November): 1–46. ISSN 00189219. doi:10.1109/5.726791.

24. Duffner, S. Face Image Analysis With Convolutional Neural Networks. *University of Freiburg Freiburg, Germany*, 2007: 191.

25. ul Hassan, M. VGG16 - Convolutional Network for Classification and Detection. *Neurohive.io*, 2018: 1–12. URL `https://neurohive.io/en/popular-networks/vgg16/`.

26. Savita N., G. and Parul, A. Detection and Classification of Plant Leaf Diseases Using Image Processing Techniques - A Review. *International Journal of Life Sciences Research*, 2014. 5(4): 120–124.

27. Dhaygude, S B., & Kumbhar, N. P. Agricultural Plant Leaf Disease Detection Using Image Processing. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2013. 2(1): 599–602. doi:ISSN:2278âĂŞ8875. URL `http://www.ijareeie.com/upload/january/5{_}Agriculturalplant.pdf`.

28. Badnakhe, M. R. and Deshmukh, P. R. An Application of K-Means Clustering and Artificial Intelligence in Pattern Recognition for Crop Diseases. *International Conference on Advancements in Information Technology With workshop of ICBMG 2011*, 2011. 20: 134–138.

29. S, A., R, N. S., S, A. and S, V. V. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International: CIGR Journal*, 2013. 15(1): 211–217. URL `http://cigrjournal.org/index.php/Ejounral/article/view/2338/1720`.

30. Kulkarni, A. H. and Ashwin, P. R. K. Applying image processing technique to detect plant diseases. *International Journal of Modern Engineering Research (IJMER)*, 2012. 2(5): 3661–3664.

31. Bashir, S. Remote Area Plant Disease Detection Using Image Processing. *IOSR Journal of Electronics and Communication Engineering*, 2012. 2(6): 31–34. ISSN 22788735. doi:10.9790/2834-0263134.

32. K.Jayamala, K. B. V. C. O. E., R.Kumar and Vidyapeeth, B. Advances in Image Processing for Detection of Plant Diseases. 2011. 2(2): 135–141.

33. Journal, I. and Patil, S. Leaf disease severity measurement using image processing LEAF DISEASE SEVERITY MEASUREMENT USING IMAGE. 2016. (January): 1–6.

34. Chaudhary, P., Chaudhari, A. K., Cheeran, A. N. and Godara, S. Color Transform Based Approach for Disease Spot Detection on Plant Leaf. *International Journal of Computer Science and Telecommunications*, 2012. 3(6): 4–9. URL `www.ijcst.org`.

35. Rathod, A. N., Tanawal, B. and Shah, V. Image Processing Techniques for Detection of Leaf Disease. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2013. 3(11): 397–399.

36. Badrinarayanan, V., Kendall, A. and Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. 39(12): 2481–2495. ISSN 1939-3539. doi: 10.1109/TPAMI.2016.2644615.