INTERNATIONAL JOURNAL OF TECHO-ENGINEERING
ISSN : 0975-4520

USING DEEP LEARNING FOR IMAGE-BASED PLANT DISEASE DETECTION

VAISHNAVI KANCHARLAPALLI[1], USHARANI PERNI[2], CHANDRIKA PUNURI[3], SHARMILA SHAIK[4], Dr. M.BHEMA LINGAIAH[5]

[1, 2, 3]UG Scholars, Department of CSE, ***MALINENI LAKSHMAIAH WOMEN'S ENGINEERING COLLEGE, GUNTUR***, India.

[4] Professor, Department of CSE, ***MALINENI LAKSHMAIAH WOMEN'S ENGINEERING COLLEGE, GUNTUR***, India.

ABSTRACT:

Rapid human population growth requires corresponding increase in food production. Easily spreadable diseases can have a strong negative impact on plant yields and even destroy whole crops. That is why early disease diagnosis and prevention are of very high importance. Traditional methods rely on lab analysis and human expertise which are usually expensive and unavailable in a large part of the undeveloped world. Since smart phones are becoming increasingly present even in the most rural areas, in recent years scientists have turned to automated image analysis as a way of identifying crop diseases. This paper presents the most recent results in this field, and a comparison of deep learning approach with the classical machine learning algorithms.

Keywords: Deep learning, ML, smart phone.

1. INTRODUCTION:

Modern technologies have given human society the ability to produce enough food to meet the demand of more than 7 billion people. However, food security remains threatened by a number of factors including climate change (Tai et al., 2014), the decline in pollinators (Report of the Plenary of the Intergovernmental Science-Policy Platform on Biodiversity

ISSN : 0975-4520

Ecosystem and Services on the work of its fourth session, 2016), plant diseases (Strange and Scott, 2005), and others. Plant diseases are not only a threat to food security at the global scale, but can also have disastrous consequences for smallholder farmers whose livelihoods depend on healthy crops. In the developing world, more than 80 percent of the agricultural production is generated by smallholder farmers (UNEP, 2013), and reports of yield loss of more than 50% due to pests and diseases are common (Harvey et al., 2014). Furthermore, the largest fraction of hungry people (50%) live in smallholder farming households (Sanchez and Swaminathan, 2005), making smallholder farmers a group that's particularly vulnerable to pathogen-derived disruptions in food supply. Historically, disease identification has been supported by agricultural extension organizations or other institutions, such as local plant clinics. In more recent times, such efforts have additionally been supported by providing information for disease diagnosis online, leveraging the increasing Internet penetration worldwide. Even more recently, tools based on mobile phones have proliferated, taking advantage of the historically unparalleled rapid uptake of mobile phone technology in all parts of the world. Smart phones in particular offer very novel approaches to help identify diseases because of their computing power, high resolution displays, and extensive built-in sets of accessories, such as advanced HD cameras. It is widely estimated that there will be between 5 and 6 billion smart phones on the globe by 2020. At the end of 2015, already 69% of the world's population had access to mobile broadband coverage, and mobile broadband penetration reached 47% in 2015, a 12-fold increase since 2007 (ITU, 2015). The combined factors of widespread smart phone penetration, HD cameras, and high

ISSN : 0975-4520

performance processors in mobile devices lead to a situation where disease diagnosis based on automated image recognition, if technically feasible, can be made available at an unprecedented scale. Here, we demonstrate the technical feasibility using a deep learning approach utilizing 54,306 images of 14 crop species with 26 diseases (or healthy) made openly available through the project Plant Village.

2 RELATED STUDY

Human population steadily continues to grow, and along with it the need for food production increases. According to the UN projections [1], human population is expected to reach 9.7 billion in 2050, 2 billion more than today. Considering that most of the population growth is to occur in the least developed countries (around 80% increase in the next 30 years), where the food scarcity is the main problem, it is easy to conclude that minimizing food loss in those countries is a primary concern. It is estimated that the yield loss worldwide is between 20 and 40 percent [2], with many farms suffering a total loss.

Traditional methods for detecting diseases require manual inspection of plants by experts. This process needs to be continuous, and can be very expensive in large farms, or even completely unavailable to many small farm holders living in rural areas. This is why many attempts to automate disease detection have been made in the last few decades. One of the notable approaches is the use of hyperspectral imaging. Hyperspectral images are usually taken by satellites or airborne imaging devices and used for monitoring large areas. A downside of this approach is extremely high equipment cost, as well as high dimensionality and small number of samples which make them unsuitable for machine learning (ML) analysis.

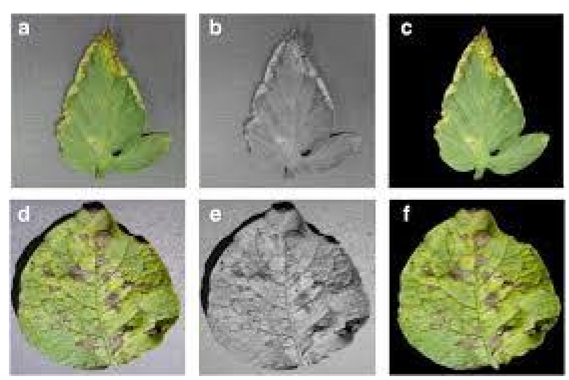
Because of the recent breakthroughs in computer vision and the availability of cheap hardware, currently the most popular approach is the analysis of RGB images. The other motive for analysing RGB images is that with the current smart phone iniquitousness these solutions have potential to reach even the most rural areas. RGB images can be analysed by classical ML algorithms or the deep learning (DL) approach. Classical methods rely on image pre-processing and the extraction of features which are then fed into one of the ML algorithms. Popular algorithm choices are Support Vector Machines (SVM), k-Nearest Neighbours (*k*-NN), Fully Connected Neural Networks (FCNN), Decision Trees, Random Forests etc. In the last few years, the researchers shifted almost exclusively to the DL methods for image classification tasks. The reason is that they almost always outperform classical algorithms when given reasonably sized dataset, and can be implemented without the need for hand-engineered features. In this paper, we compare the DL approach with classical ML algorithms for the study case of plant disease classification.

3. PROPOSED SYSTEM:

Crop diseases are a major threat to food security, but their rapid identification remains difficult in many parts of the world due to the lack of the necessary infrastructure. The combination of increasing global smart phone penetration and recent advances in computer vision made possible by deep learning has paved the way for smart phone-assisted disease diagnosis. Using a public dataset of 54,306 images of diseased and healthy plant leaves collected under controlled conditions, we train a deep convolutional neural network to identify 14 crop species and 26 diseases (or absence thereof). The trained model achieves an accuracy of 99.35% on a held-out test set, demonstrating the feasibility of this approach. Overall, the approach of training deep learning models on increasingly large and publicly available image datasets presents a clear path toward smart phone-assisted crop disease diagnosis on a massive global scale.

RESULTS:

SVM is a supervised learning algorithm used for classification or regression problems. Classification is done by defining a separating hyperplane in the feature space. In the original form, it performs linear classification on two classes. By using kernels, it can also perform nonlinear classification. Kernels are used for an efficient transformation of the original feature space into high dimensional or infinite dimensional feature space, allowing for highly non-linear hyperplanes. SVM can fit highly complex datasets and at the same time exhibit good generalization properties. Multiclass classification using SVM can be implemented using one-vs-all or one-vs-one strategies. One-vs-all approach trains N classifiers (N is the number of classes), where each classifier considers examples from one class as positive, and all others as negative. One-vsone approach trains N(N-1)/2 binary classifiers and determines the winner by max-wins voting [10]. We have experimented with different configurations and found that the best results were obtained by using the radial basis function kernel and regularization parameter C=100. One-vs-all approach was used. The achieved accuracy on the test set is 91.74%.



4. CONCLUSION:

This paper presents the dominance of the DL method over the classical ML algorithms. Both the simplicity of the approach and the achieved accuracy confirm that the DL is the way to follow for image classification problems with relatively large datasets. As the achieved accuracy of the DL method is already very high, trying to improve its results on the same dataset would be of little benefit. Further work with the DL model could be done by expanding the dataset with more diverse images, collected from multiple sources, in order to allow it to generalize better. The considered ML algorithms achieved relatively high accuracy, but with error rates still an order of magnitude higher than the DL model. Further work in improving accuracy of the classical approach can be done by experimenting with other algorithms and by improving the features, as most likely they are the limiting factor of this approach.

REFERENCES:

[1] United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019: Highlights (ST/ESA/SER.A/423).

[2] Savary, Serge, et al. "The global burden of pathogens and pests on major food crops." *Nature ecology & evolution* 3.3 (2019): 430.

[3] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7 (2016): 1419.

[4] Fujita, E., et al. "A practical plant diagnosis system for field leaf images and feature visualization." *International Journal of Engineering & Technology* 7.4.11 (2018): 49-54.

[5] Haralick, Robert M., Karthikeyan Shanmugam, and Its' Hak Dinstein. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 6 (1973): 610-621.

[6] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20.3 (1995): 273-297.

ISSN : 0975-4520

[7] Cunningham, Padraig, and Sarah Jane Delany. "k-Nearest neighbour classifiers." Multiple Classifier Systems 34.8 (2007): 1-17.

[8] Haykin, Simon. Neural networks: a comprehensive foundation. Prentice Hall PTR, 1994.

[9] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[10] Duan, Kai-Bo, and S. Sathiya Keerthi. "Which is the best multiclass SVM method? An empirical study." International workshop on multiple classifier systems. Springer, Berlin, Heidelberg, 2005.

[11] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.