Multimodal Transfer Learning Outlook for Plants Disease Prognosis

Anshul Dhingra, Samarth Khajuria

Delhi Technological University

---***---

Abstract

Crop diseases are a significant problem on the road to food hygiene and security, but still they cannot be identified rapidly in many parts of the world due to the lack of essential infrastructure. But Due to increasing global smartphone penetration and recent advances in Artificial Intelligence and Computer Vision has made it possible by using deep learning and has paved the way for device-assisted disease diagnosis. By Using a dataset of approximately 54,000 images of diseased as well as healthy plant leaves gathered under controlled conditions, we train a deep convolutional neural network for the identification of 14 crops and their 26 diseases. After the training phase, the model achieves an accuracy of 99.06% on a test set. Overall, by using the method of training deep learning models on large and increasing publicly available image datasets presents a path towards device-assisted crop disease prognosis on a vast global scale.

1 Introduction

In the 21st Century, Latest Technologies have enabled human society to produce food to meet the demand of more than 6 billion people. But, food hygiene and security remains a threat by a number of reasons including climatic change, plant diseases etc. Plant diseases not only creates a threatening scenario to food hygiene at the global level but can also create damaging consequences for small-scale farmers whose lives depend on crops. In this flourishing world, more than 70 percent of the agro-production is generated by small scale farmers, and reports of loss of more than 60% due to diseases are there. Moreover, the hungry people accounting for more than 40% live under small scale farming households, affecting the vulnerability of this group to pathogen-derived breakage in the food chain.

Various Efforts have been made to prevent crops from diseases. Going back in history, the application of pesticide in the past have been transformed by IPM (Integrated Pest Management) approaches. Disease identification has been largely guided by the extension of agricultural organisations, such as local clinics related to plantation and agriculture. In recent times, efforts have been guided by transferring information and knowledge for disease prognosis online, averaging the increase of web penetration. Personal Computers in particular offer very unique approaches for the identification of disease because of their super high computational power, High-Resolution displays and accessories including high definition cameras. At the end of 2019, already most of the population of the world have access to computers and the internet, a 20-fold increase since 2005. Here, we explain the technological feasibility of using a Deep Learning CNN approach with transfer learning demonstrating by applying on a dataset of approximately 54,000 images of 14 crops with 26 diseases made publicly available by Plant Village project.

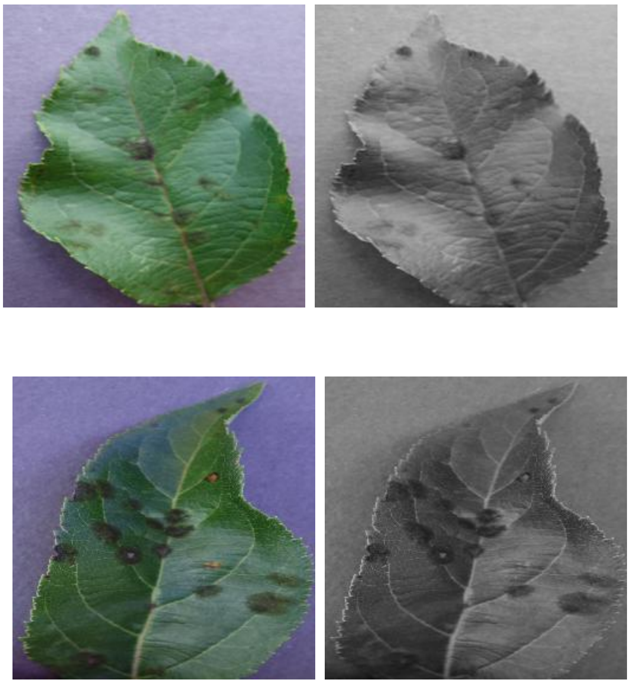
Deep Learning and object recognition, has made huge advances in the past years. The PAS-

CAL VOC Challenge and ILSVRC (Large Scale Visual Recognition Challenge) which were based on ImageNet dataset have been regarded as a criterion for wide visualisation-related obstacles in Computer Vision, including object detection. Recently, a deep CNN network achieved a top 5 error of nearly 16% for the image classification into 1000 classes, but now the error has been reduced to 3.57% in these 3 years. Moreover, training large deep neural networks can be very time consuming, the pre-trained models can classify images very quickly, making them more convenient for applications on personal computers.

Deep Neural Networks have been recently applied successfully in many different fields as end-to-end learning models. Neural networks comprise basically a mapping between an input is an image of a crop which is diseased as in our model to an output which is a crop-disease pair in our model. A neural network comprises of many hidden layers further consisting of nodes or neurons which are mathematical functions that take inputs (mathematical) from the adjoining edges and will further provide a mathematical output via an output edge. These Neural networks basically maps the input layers to the output softmax layers via various hidden layers stacked in between. The aim and goal is to develop a deep network in a way that all the neurons (nodes) weigh correctly and possesses the correct Mathematical function and same goes for connecting edges. Deep networks are trained by tuning the weights also known as network parameters in a way in which mapping improves during the training phase, back-propagating the loss by comparing the predicted and actual output further back-propagating the loss.

To develop and increase the accuracy measure of image classifiers for plant disease detection and prognosis we needed a larger, verified dataset containing images of diseased as well as healthy plants. So, we used Plant Village project which is collecting hundreds of thousands of plant images and is further increasing day by day furthermore being freely available to the public on GitHub. Here, we demonstrate the classification of 26 diseases among 14 different crops using approximately 54000 images with a CNN based transfer learning approach. We calculate the performance of our model based on its ability to predict correct crop, as well as disease in pair among, made 38 classes in pairs. Our best-trained model achieves a mean F1 score of 0.9906 (accuracy of 99.06%), our results giving a step forward towards a computer-assisted plant disease prognosis system.

2 Dataset Information



We analyse approximately 54,000 images of disease plants, having a 38 category labels. Each category is a crop-disease in pair, and we try to guess the same, given just the image of the diseased plant. (Image. 1) shows each example of crop-disease pair from the Plant Village project dataset. We have further resized the image to 256X256 pixels, performing different feature extraction & Deep Learning techniques on these downscaled images.

Among all the models, we used 2 different versions of the same dataset Plant Village project. 2 versions were (a) Coloured Images, (b) Grey-scaled Images. We started 1st with colour images, then grey-scaled. This set is constructed to understand if the CNN model is actually learning the "impression" of the plant diseases, or is it just learning the "constitutional bigotry" in the dataset. (Image 2) shows the different versions of the same plant.

3 Performance Check

To get an estimate how our constructed model works, and also to know if our approaches are overfitting, we run all our model experiments across a range of train-test splits, for example 70–30 (70% of the whole dataset used for training, and 30% for testing), 80–20 (80% of the whole dataset used for training, and 20% for testing), 60–40 (60% of the whole dataset used for training, and 40% for testing), 50–50 (50% of the whole dataset used for training, and 50% for testing) and at last 20–80 (20% of complete dataset used for training, and 80% of the whole dataset for testing). We observe, the Plant Village Project dataset has images of the same plant leaf (taken from different angles), and we tried to do the mappings of such cases for 40,000 images out of the 54,000 images; and during running of all these test-train splits, we make sure all the images of the same plant disease inputs either in training set or test set. Moreover, after running every model experiment, we measure the mean precision, recall, & F1 score, along with the aggregate accuracy during the complete epoch period of training(at the end of every epoch). We use the final mean F1 score for the comparison and calculation of results among all the different experimental versions and configurations.

4 Approach

We examine the applicability of Deep CNN networks for the classification task. We focus on two in-demand architectures, namely ResNet(50) and GoogLeNet which were developed in ILSVRC (Large Scale Visual Recognition Challenge) for the ImageNet dataset. At the ILSVRC 2015, The Residual Neural Network (ResNet) by Kaiming He et al developed an architecture with "skip connections" which featured heavy batch normalization. Skip connections are known as gated units or gated recurrent units in fact they have a strong similarity to successful elements applied in RNNs. It achieves a top-5 error rate of 3.57ResNet architecture is comprised of VGG-19 at the bottom (A state-of-the-art approach in Large Scale Visual Recognition Challenge 2014), then 34-layer plain network is treated as the deeper network of VGG-19, i.e more convolution layers, and then further 34=layer residual network at the top which is a plain one with skip connection. All the layers usually have ReLu non-linear activation units associated with them. At the last, fully connected layer has 38 outputs in our version of ResNet(which equals to the total number of categories or classes which is also 38 in our dataset), which further feeds to the SoftMax layer. The idea of using a ResNet is through one or more layers, whereas in the deep CNNs in general all layers are processed one at a time. The GoogLeNet architecture is a deep and wide architecture with 22 layers in general, but still with a lower number of parameters (just approximately 5 million) whereas there are more than 23 million in ResNet-50. GoogLeNet works on the basic principle of 'Network in Network' architecture in the form of inception models. It uses a parallel 1X1, 3X3 and 5X5 convolutions across max-pooling layers in parallel. This is the way a single inception model is formed, whereas we will be using 9 modules similar to this for our experiment. We examine the performance of these both models that is ResNet and GoogLeNet on the Plant Village dataset by training both of them using the transfer learning approach. In transfer learning, we reinitialise the weights of fully connected layers in ResNet and of loss (1,2,3) layers in case of GoogleNet. To compile, we have a total of 20 experimental configurations, depending on the following parameters: begin 1. Choice of Transfer Learning Architecture: ResNet GoogLeNet 2. Choice of the dataset to be used: Colour, Gray-Scaled 3. Training-Test Split:

- Train: 80% Test: 20%
- Train: 80% Test: 20%
- Train: 80% Test: 20%
- Train: 70% Test: 30%
- Train: 60% Test: 40%
- Train: 50% Test: 50%

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 01 | Jan 2021 www.irjet.net p-ISSN: 2395-0072

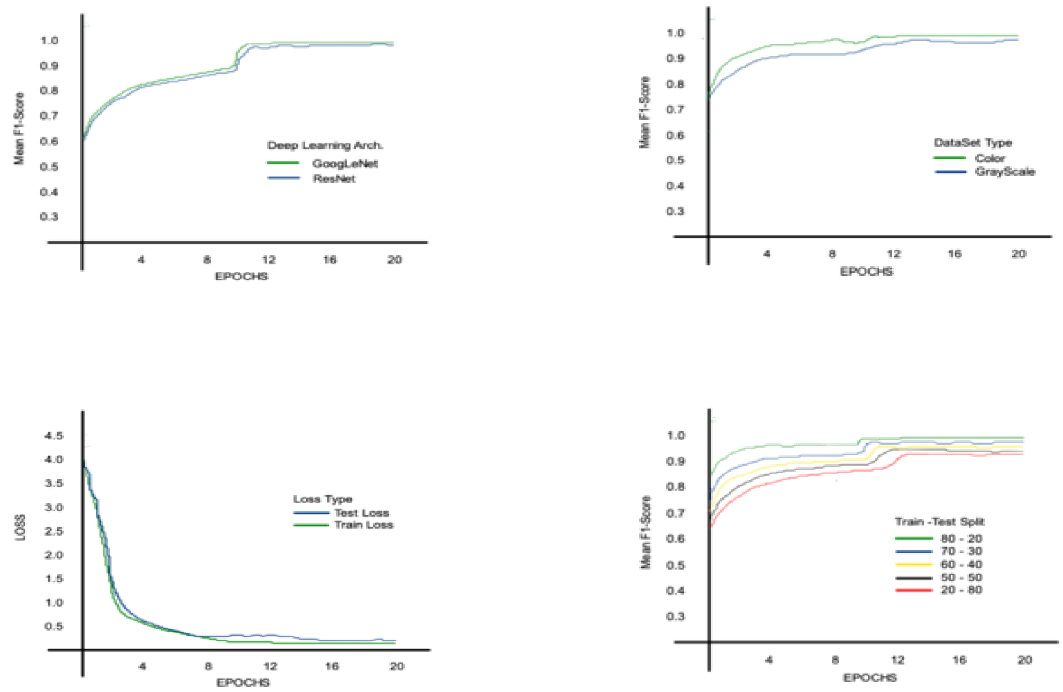
- Train: 20% Test: 80%

Each of these 20 experiments runs for a total of 20 epochs, where one epoch is the number of training repetitions in which our Neural Network completes one cycle of whole training set. For a fair comparison and observation, we standardize the parameters also known as hyper-parameters in all the 20 experiments:

1. Optimizer: Stochastic Gradient Descent,
2. Learning Rate: 0.004
3. Learning Rate Variation: Step (Decrease of 10 after every 20/2 epochs)
4. Momentum: 0.9
5. Weight Degenerate: 0.0005,
6. Batch Size: 50(Resnet), 16(GoogLeNet)
7. Gamma: 0.1

All the above experiments were performed using Keras with TensorFlow backend which is a fast, open-source framework for deep learning and achieves good results, especially on the media data.

5 Results



Graphs showing the various properties of our best perofrmed experimental version.

We observed and calculated that on a dataset with 38 classes/categories, a random guess will achieve nearly 2.5% accuracy on average. Across all of our experiments on the Plant Village Dataset, the overall accuracy we obtained varied from 79.9% (in case of ResNet::GrayScale::20-80) to 99.06% (in case of GoogLeNet::Coloured::80-20), hence it shows a strong range of accuracy variation. Further to address the problem of overfitting, we changed the training to test set ratio and observed that even for the last case that is on 20% training data and rest 80% testing data, model achieved a satisfactory 96.7% accuracy(GoogLeNet::Color::20-80). But we did saw that accuracy decrease if we keep decreasing the share of the dataset for the training set.

TRANSFER LEARNING

	RESNET	GOOGLENET
TRAIN 20% TEST 80%		
COLOUR	0.9356	0.9679
GREYSCALE	0.7941	0.8321
TRAIN 50% TEST 50%		
COLOUR	0.9566	0.9772
GREYSCALE	0.8561	0.9112
TRAIN 60% TEST 40%		
COLOUR	0.9698	0.8913
GREYSCALE	0.9849	0.9348
TRAIN 70% TEST 30%		
COLOUR	0.9531	0.9901
GREYSCALE	0.9094	0.9448
TRAIN 80% TEST 20%		
COLOUR	0.9665	0.9906
GREYSCALE	0.9394	0.9499

The models performed better on the coloured datasets than grayscale datasets. We were concerned that the CNN network might pick up the inherent biases only, therefore we experimented with the grey-scale dataset to test the ability of the model to recognize the disease and the crop in absence of colour information. As a result, the performance did degrade in absence of colour information, but even in the worst-case scenario the accuracy was quite satisfied that is 79%. To summarize, all the results have been concluded with model recognizing the crop and the disease both simultaneously whereas we believe if the dataset could be modified with the type of crop already mapped, we could yield much better results.

6 Discussion

The performance of Deep Learning & Convolutional Neural Networks in image classification and object recognition has made huge progress in recent years. In the past, the traditional methods for image classification tasks were based on hand-designed or engineered features, such as SIFT, HoG, SURF, etc., and then people used some form of a learning algorithm in these feature matrices. The performance of these methods therefore relied heavily on the associated predefined features. Feature Engineering was a very tedious and complex process which needed to be revisited time to time when the problem in hand or the associated dataset changes significantly. This problem occurs in with all the usual methods to detect plant diseases using computer vision as they depend mainly on hand-engineered features, image enhancement techniques, and a lot of other complex and labour-intensive technologies. Our method is based on the work which showed that end-to-end supervised training using transfer learning on a deep CNN network architecture is a practical approach even for image classification and object detection problems having a very large number of categories or classes, getting ahead of the traditional methods using hand-engineered features. Skipping the labour-intensive task of feature engineering and the familiarity of the solution makes them a very encouraging candidate for a constructive and expandable path for computational interpretation of crops and plant diseases. The limitation is that we are at the moment compulsive to the analysis of single leaves, facing up, on a more or less same background. As this is not in the case of real-life leaves images, real-life analysis and prognosis should be able to classify and analyse images of a disease as it presents itself directly on the plant. In fact, most of the diseases don't present themselves on the upper or lower side of leaves only, but on different parts of the crop species. Thus, we would try to collect new image data with more realistic images and most importantly from many different perspectives. By using 38 categories that comprise of both crop species and disease type, we have even made the task harder than necessary from a practical view, as people related to agricultural and farming are expected to know which crop species they are growing and harvesting. Be it the high accuracy on the Plant Village Project Dataset, if we limit the classification challenge to the disease status then it won't have a measurable effect. But, on the existing real-world datasets, we can conclude significant improvements inaccuracy. Overall, the experimented method works reasonably well with any type of distinguishable crops and diseases, and further to improve quite noticeably with more training data. Decisively, it's worth to note that the method presented here does not aim to replace the already existing solutions for disease prognosis, but rather to fortify them. As we all know, Laboratory tests are always more reliable than prognosis based on just visual symptoms, and even most of the times early-stage prognosis via visual inspection alone is quite difficult. Coming to the point, there are more than 2.5 billion personal computers in the world by 2020—of which almost 20 million alone in Africa, we do think that this method represents a viable additional method to help decrease the yield loss. Strongly, in the future, image data from a personal computer may be fortified with location and as well as time information for additional improvements in the measurement of the accuracy and precision. Lastly, we do have our belief on the astonishing pace at which computer technology has developed in recent years, and will continue to do so. Also with improving number and quality of processing power of computing devices, we consider it likely that highly accurate prognosis via computers is only a question of time.

References

[1] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature 521, 436–444. doi: 10.1038/nature14539

[2] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, eds F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc.),

1097–1105.

[3] Mokhtar, U., Ali, M. A., Hassanien, A. E., and Hefny, H. (2015). “Identifying two of tomatoes leaf viruses using support vector machine,” in Information Systems Design and Intelligent Applications, eds J. K. Mandal, S. C. Satapathy, M. K. Sanyal, P. P. Sarkar, A. Mukhopadhyay (Springer), 771–782.

[4] Zeiler, M. D., and Fergus, R. (2014). “Visualizing and understanding convolutional networks,” in Computer Vision–ECCV 2014, eds D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Springer), 818–833.

[5] Strange, R. N., and Scott, P. R. (2005). Plant disease: a threat to global food security. Phytopathology 43, 83–116. doi: 10.1146/annurev.phyto.43.113004.133839

[6] Desai M, Jain AK, Jain NK, Jethwa K. Detection and classification of fruit disease: a review. Int Res J Eng Technol 2016;3(3):727–9.

[7] Deng J, Dong W, Socher R, Li L-J, Li K, Li F-F. 2009. ImageNet: a large-scale hierarchical image database. In: IEEE conference on computer vision and pattern recognition. 248–255 DOI 10.1109/CVPRW.2009.5206848.

[8] ”PlantVillage”, Plantvillage.psu.edu, 2020. Available: https://plantvillage.psu.edu/.

[9] D. Klauser, ”Challenges in monitoring and managing plant diseases in developing countries”, Journal of Plant Diseases and Protection, vol. 125, no. 3, pp. 235-237, 2018. Available: 10.1007/s41348-018-0145-9.

[10] D. Wu, Y. Wang, S. Xia, J. Bailey and X. Ma, ”Skip Connections Matter: On the Transferability of Adversarial Examples Generated with ResNets”, arXiv.org, 2020. Available: https://arxiv.org/abs/2002.05990.

[11] P. Sharma, Y. Berwal and W. Ghai, ”Performance analysis of deep learning CNN models for disease detection in plants using image segmentation”, Information Processing in Agriculture, 2019. Available: 10.1016/j.inpa.2019.11.001.

[12] 12. S. Mohanty, D. Hughes and M. Salathé, ”Using Deep Learning for Image-Based Plant Disease Detection”, Frontiers in Plant Science, vol. 7, 2016. Available: 10.3389/fpls.2016.01419.

[13] Y. Oo and N. Htun, ”Plant Leaf Disease Detection and Classification using Image Processing”, International Journal of Research and Engineering, vol. 5, no. 9, pp. 516-523, 2018. Available: 10.21276/ijre.2018.5.9.4

[14] S. Mohanty, ”spMohanty/PlantVillage-Dataset”, GitHub, 2016. [Online]. Available: https://github.com/spMohanty/PlantVillage-Dataset/tree/master/raw/color.

[15] Atabay, H. A. 2016b. A convolutional neural network with a new architecture applied on leaf classification. IIOAB J 7(5):226–331

[16] P. Pawara, E. Okafor, L. Schomaker and M. Wiering, ”Data Augmentation for Plant Classification”, 2017.

[17] Isleib, ”Signs and symptoms of plant disease: Is it fungal, viral or bacterial?”, MSU Extension, 2012.

[18] S. Sankaran, A. Mishra, R. Ehsani and C. Davis, ”A review of advanced techniques for detecting plant diseases”, Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1-13, 2010. Available: 10.1016/j.compag.2010.02.007.

[19] 19. ”Practical Deep Learning for Coders, v3 — fast.ai course v3”, Course.fast.ai, 2020.

[20] ”fastai”, Docs.fast.ai, 2020. [Online]. Available: https://docs.fast.ai/.