# LITERATURE REVIEW

A PREPRINT

**Huilin Chang**
School of Data Science
University of Virginia
Charlottesville, VA 22904
hc5hq@virginia.edu

**Yihnew Eshetu**
School of Data Science
University of Virginia
Charlottesville, VA 22904
yte9pc@virginia.edu

**Celeste Lemrow**
School of Data Science
University of Virginia
Charlottesville, VA 22904
ctl7t@virginia.edu

February 26, 2021

## Overview

As both global population growth and effects of climate change increase, the field of agriculture research faces an increasing need for analytic approaches that support increased yields, improved planting and harvesting efficiency, and mitigation of pests and diseases that negatively impact food production across all regions of the world. In addition, with the increasing globalization of food production and food distribution, there is a need for greater technological efficiency in the monitoring and study of crop production to provide more direct and comprehensive support to farmers in the field at a faster pace and lower labor cost. Several of these identified research needs align with the comparative advantages of deep learning. Opportunities to leverage computer vision to study crop yields and diseases and pests, quickly synthesize larger datasets of satellite imagery to assess plant growth, and predict production volume and potential environmental condition impacts to production all can lend themselves to deep learning techniques. This literature review presents a brief synthesis of a corpus of papers we curated that focus on the use of deep learning for precision agriculture, pest and disease identification, yield prediction, and environmental prediction, in support of our intended research focus on deep learning for agricultural applications

## Deep Learning in Agriculture

Overall, the set of references we surveyed demonstrated successful implementations of deep learning techniques, including: focusing on image classification related to positive or negative pest and disease identification [7, 11, 13] in support of maximizing crop production; prediction of environmental conditions to support decision-making about planting and cultivation technique adjustments [2], and computer vision to support the use of robotics to automate agricultural labor tasks, such as harvesting or weed removal [6]. Some studies relied exclusively on deep learning algorithms for classification [3]; others utilized a hybrid approach, turning to deep learning for advanced feature extraction and then using a supervised machine learning algorithm, such as support vector machines, for the classification phase [11]. The surveyed studies provided useful insights to support initial parameter selection and optimization approaches for models, specifying details such as activation function, number of layers, and use of pooling layers [1, 3], which can be leveraged in the design and adjustment of models in future studies. Many of the studies we examined mention the importance of resizing images to impact the model. From one study to another, we noticed unique techniques were used for image manipulation. Furthermore, certain studies built their own neural networks while others used existing architecture such as LeNet as a foundation for their network.

## Data

Several of the studies discussed relevant gaps and considerations specific to applying deep learning to the agriculture sectors. One issue of note, for example, is the lack of datasets of sufficient size, labeling, and image quality; larger and more diverse image sets are needed to support the range of potential research studies [6]. In the case of lack of data many studies used data augmentation techniques such as scaling, cropping, and brightening to increase their dataset

[12] Useful references to available datasets, however, were provided and discussed in one study, which we plan to use as part of our dataset review and selection [6].

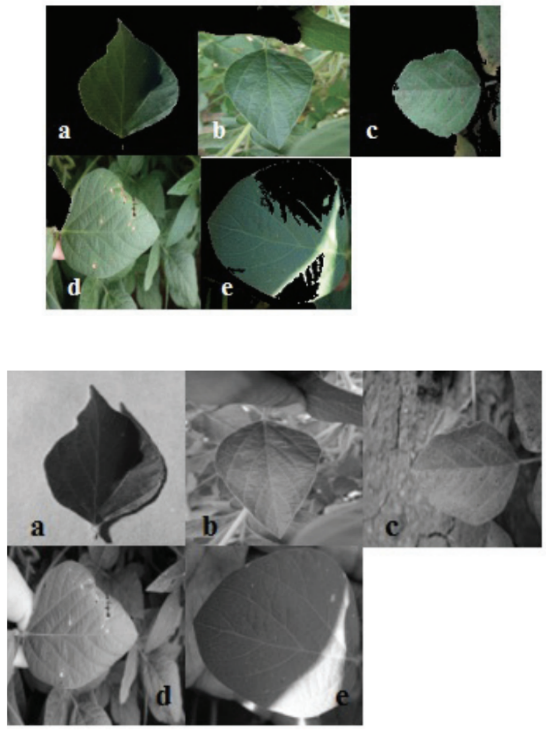


Figure 1: (**Top**): Resized Color Images (**Bottom**): Resized Grayscale Images [12]

One study explores the impact resizing has on images that are originally of different resolution. As seen in Figure 1, in order to prevent any bias, this particular study decided to resize all images to 128x128 pixels [12]. Furthermore, since many of the images were taken under different lighting conditions, they decided to create a separate dataset containing grayscale images to reduce noise, Figure 1. Using the two datasets, the test classification accuracy was compared.

## Data Pipeline

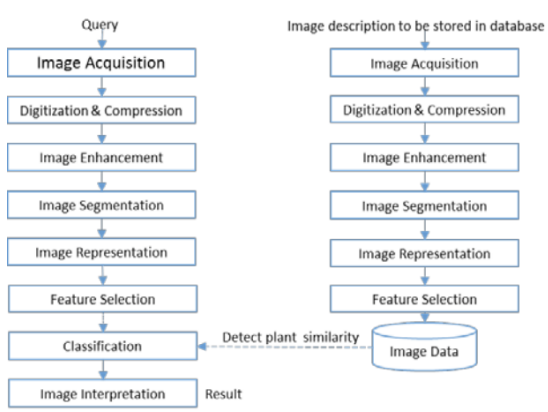


Figure 2: Detailed Illustration of an Image Pipeline [1]

In addition, we noted some useful discussion about data pipeline development to consider when dealing with satellite imagery, in particular, such as spectral normalization, solar correction, and atmospheric correction. We also found useful assessment of feature engineering approaches for different types of image information, such as spectral, spatial, and temporal information [8]. While these studies provided useful initial insights we can leverage in both data selection and data modeling and analysis, they also highlighted that there is room for growth in understanding about how these considerations can be studied further to enhance these types of deep learning models.

Figure 2, depicts a detailed view of the data pipeline architecture used in one study [1]. In this pipeline we see steps such as feature extraction, data augmentation, and image analysis.

## Image Processing

Furthermore, there seem to be gaps in the literature regarding broader issues of image quality and image processing and engineering, and how those parameters impact both data processing decisions and performance and quality of models. Further investigation on the impact on analysis of different types of images, such as close-up photos compared with satellite imagery, may be warranted. While some papers discussed the merits and concerns of a particular type of imagery of interest, we did not find a lot of comparative discussion across image types, and how that may impact the development of a model and its results. This may be due to the particular set of references we have reviewed so far for this project; we will continue to look at this as part of additional background research.

In the specific area of deep learning for agricultural disease recognition, we found discussion of some particular limitations, including consideration of whether there is a large enough labeled dataset, which is important to prevent or limit overfitting. In addition, more complex models may rely such a large number of features and become so specialized to a particular disease classification or dataset, it’s external validity may be very limited, which, in turn, raises the issue of high computational costs and effort to train new models from scratch for multiple contexts[13].

## Conclusion

Overall, this literature review proved useful in helping us identify and formulate initial research questions for consideration as we move forward with dataset selection and full proposal development for our project. Based on the key themes and findings from the work we reviewed, some potential questions include:

### Research Questions

- How can we use automation and artificial intelligence methods to solve traditional farming problems, given agriculture’s significant role in the global economy?
- How can we leverage existing deep learning models and architecture to help solve issues in agriculture?
- Based on current research, are there areas of agricultural research where deep learning approaches have not been implemented, but have potential to be used given potential common characteristics and linkages with existing studies?
- Would the merging of datasets possibly help improve certain deep learning agriculture models? If so, what approaches would be best for doing that?
- How can we improve upon the current limits of highly-specialized models and limited external validity in studies focusing on deep learning for pest and disease identification to support more broadly applicable research?

As we move forward with data selection and research question development, and consider which sub-area to focus on, our assessment is that there is currently more completed research work regarding identification of plant disease and pests, although we noted that there seems to be more of a challenge in moving beyond just classification of a diseased plant, and determining the specific type or cause of the disease infestation. Work related to determination of yield, prediction of yield, and identification and prediction of environmental conditions appears to be more sparse. In addition, linked to the paucity of available image data, there is also space to generate classification models for additional species and types of crops. A key next step will be honing in further on a dataset that meets necessary criteria for deep learning in conjunction with determination of a more specific research focus that supports a Virginia-based need. Based on our review of datasets so far, we have identified several promising image collections with linkages to prominent Virginia crops, such as apples, that would provide utility for state-focused precision agriculture issues.

## References


[1] Abdullahi, Halimatu Sadiyah, and Ray E. Sheriff. "Convolution Neural Network in Precision Agriculture for Plant Image Recognition and Classification." *International Conference on Innovative Computing Technology 2021. IEEE Xplore.* Web. 26 Feb. 2021.

[2] Chen, Shuchang, Bingchan Li, Jie Cao, and Bo Mao. "Research on Agricultural Environment Prediction Based on Deep Learning." *Procedia Computer Science.* Elsevier, 18 Oct. 2018. Web. 26 Feb. 2021.

[3] Dahikar, Snehal and Rode, Sandeep. "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach." *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering.* January 2014. Web. 26 Feb. 2021.

[4] Espejo-Garcia, Borja, Nikolaos Mylonas, Loukas Athanasakos, and Spyros Fountas. "Improving Weeds Identification with a Repository of Agricultural Pre-trained Deep Neural Networks." *Computers and Electronics in Agriculture.* Elsevier, 29 June 2020. Web. 26 Feb. 2021.

[5] Khan, Tamoor, Jiangtao Qiu, Muhammad Asim Ali Qureshi, Muhammad Shahid Iqbal, Rashid Mehmood, and Waqar Hussain. "Agricultural Fruit Prediction Using Deep Neural Networks." *Procedia Computer Science.* Elsevier, 27 July 2020. Web. 26 Feb. 2021.

[6] Lu, Yuzhen, and Sierra Young. "A Survey of Public Datasets for Computer Vision Tasks in Precision Agriculture." *Computers and Electronics in Agriculture.* Elsevier, 08 Sept. 2020. Web. 26 Feb. 2021.

[7] Nasir, Inzamam Mashood, Bibi, Asima, Shah, Jamal Hussain, Khan, Muhammad Attique. "Deep Learning-Based Classification of Fruit Diseases: An Application for Precision Agriculture." *Computers, Materials & Continua.* Web. 26 Feb. 2021.

[8] Nguyen, Thanh Tam, Thanh Dat Hoang, Minh Tam Pham, Tuyet Trinh Vu, Thanh Hung Nguyen, Quyet-Thang Huynh, and Jun Jo. "Monitoring Agriculture Areas with Satellite Images and Deep Learning." *Applied Soft Computing.* Elsevier, 23 July 2020. Web. 26 Feb. 2021.

[9] Sladojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification." *Computational Intelligence and Neuroscience.* Hindawi, 22 June 2016. Web. 26 Feb. 2021.

[10] Tahir, Muqadas Bin, Muhammad Attique Khan, Kashif Javed, Seifedine Kadry, Yu-Dong Zhang, Tallha Akram, and Muhammad Nazir. "Recognition of Apple Leaf Diseases Using Deep Learning and Variances-Controlled Features Reduction." *Microprocessors and Microsystems.* Elsevier, 15 Jan. 2021. Web. 26 Feb. 2021.

[11] Tombe, Ronald. "Computer Vision for Smart Farming and Sustainable Agriculture." *2020 IST-Africa Conference (IST-Africa) IEEE Xplore.* Web. 26 Feb. 2021.

[12] Wallelign, Serawork, Mihai Polceanu, and Cédric Buche. "Soybean Plant Disease Identification Using Convolutional Neural Network." *Artificial Intelligence Research Society Conference.* 05 June 2018. Web. 26 Feb. 2021.

[13] Yuan, Yuan, Lei Chen, Huarui Wu, and Lin Li. "Advanced Agricultural Disease Image Recognition Technologies: A Review." *Information Processing in Agriculture.* Elsevier, 30 Jan. 2021. Web. 26 Feb. 2021.

[14] Zapotezny-Anderson, Paul, and Chris Lehnert. "Towards Active Robotic Vision in Agriculture: A Deep Learning Approach to Visual Servoing in Occluded and Unstructured Protected Cropping Environments." *IFAC-PapersOnLine. Elsevier, 31 Dec. 2019.* Web. 26 Feb. 2021.

[15] Zheng, Yang-Yang, Jian-Lei Kong, Xue-Bo Jin, Xiao-Yi Wang, Ting-Li Su, and Min Zuo. "CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in Precision Agriculture." *MDPI. Multidisciplinary Digital Publishing Institute*, 01 Mar. 2019. Web.