# Residual Neural Networks for Waterlogged Field Classification

**Trevor Wong, Yaping Cai, Kaiyu Guan**
**Department of Natural Resources and Environmental Sciences, College of Agricultural, Consumer, and Environmental Sciences, University of Illinois at Urbana-Champaign**

## INTRODUCTION

### What is Waterlogging?

After large amounts of precipitation or overirrigation, the ground in fields may become saturated with water and form pools. This phenomenon, known as waterlogging, can have adverse effects on fields, including[1]:

- Poor aeration for plant roots
- Loss of soil structure
- Soil pH and salinity becoming unsuitable for crop productivity
- Soil nutrient imbalances
- Difficulty cultivating waterlogged crops

*Fig. 1: Waterlogged soybean field[3]*

## OBJECTIVES

The objective of this project is to enable the classification of waterlogged fields using satellite imagery.

### Scalability

One goal of this project is to enable high scalability, such that it can be applied to many locations with satellite data and field bounds without the need to collect data on-location. Even where field bounds are unavailable, they may be possible to extrapolate using information from the Crop Data Layer provided publicly by the USDA. This would allow data to be collected for a large geographic area and over a long period of time.

### Generality

Another goal of this project is to ensure that it is applicable to a wide variety of locations and crops. This method does not make assumptions about what crop is being grown or which farming techniques are being used. Given a field is large and unobstructed by clouds or shadows, satellite imagery should contain enough information to determine the extent to which a field is waterlogged.

## METHODS

### Data Preprocessing

Training a model to classify waterlogged fields requires many samples of waterlogged and non-waterlogged fields. We used a satellite image of Illinois with a 3-meter resolution and field bounds derived from the publicly available Crop Data Layer to generate field-level images. Next, we sorted these images into fields that lacked waterlogging and fields that were significantly waterlogged. Finally, augmented the data with reflections and rotations to increase the amount of training data eight-fold. Figures 2, 3, and 4 show some examples of fields used as input data.

### Neural Network Architecture

To classify fields, we used an architecture known as Resnet-50. Deeper neural networks tend to be better at image recognition, but they also suffer issues with overfitting and higher training error. Residual neural network address these shortcomings with "shortcut connections," which directly map shallow layers to deeper layers[2].

### Training and Evaluation

We trained our model while varying hyperparameters such as batch size and learning rate with a fixed number of epochs. We randomly split the dataset, using 80% for training and 20% for validation.

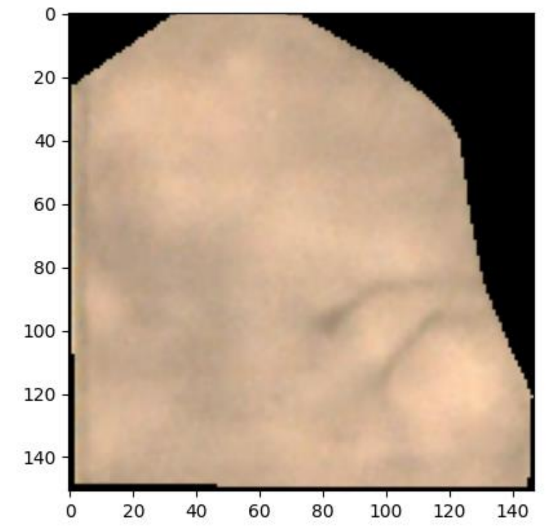

*Fig. 2: Sample field from Illinois that does not contain significant waterlogging. While there are some variations in color, there are no round, dark patches that would indicate a waterlogged area.*

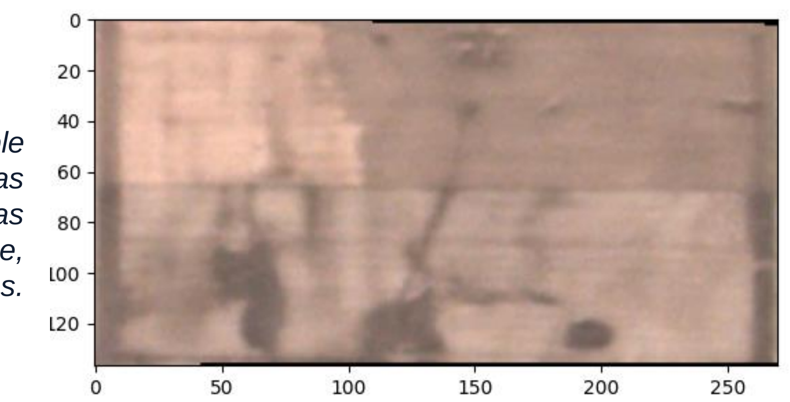

*Fig. 3: Another sample field from Illinois has severe waterlogging, as shown in the large, round dark patches.*

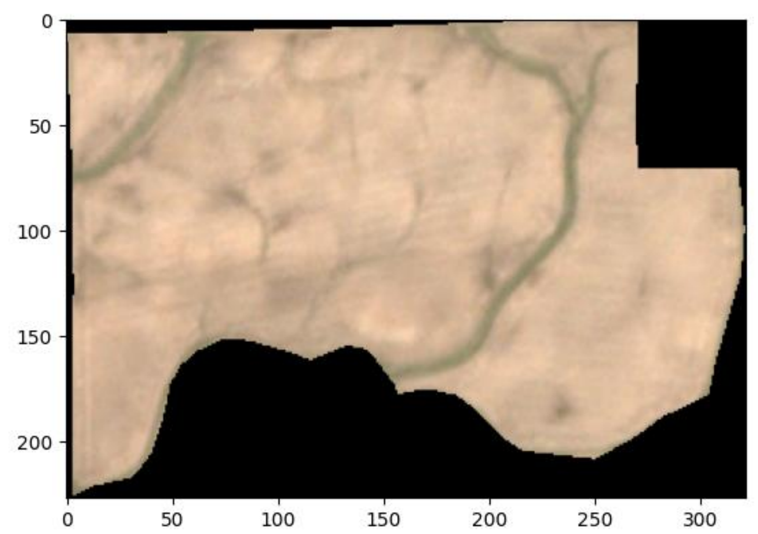

*Fig. 4: Not all fields with water are waterlogged. For instance, this field has waterways that are often used to reduce waterlogging*

## RESULTS

### Model Training

While training the model to classify fields, we tracked the accuracy, loss, validation accuracy, and validation loss for various hyperparameters. Some of these results are plotted in figures 5 and 6, shown below.

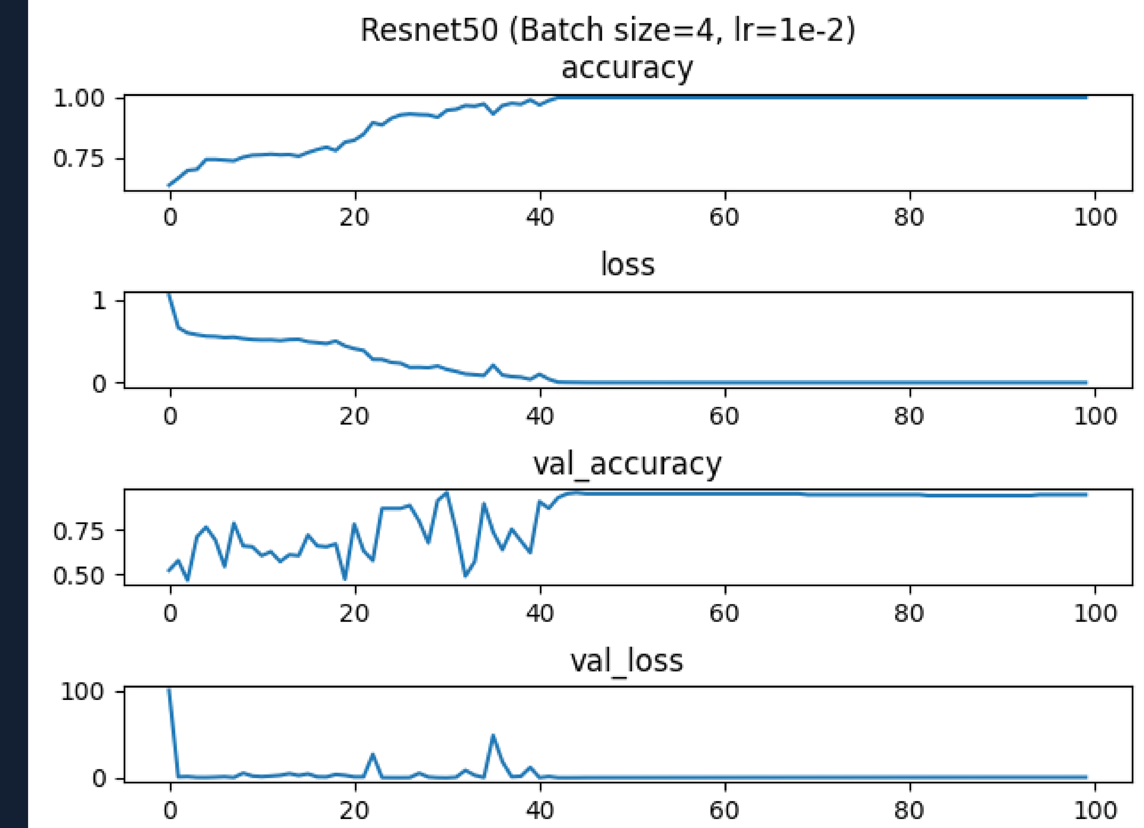


*Fig. 5: In these training graphs, there is a great deal of instability in the validation accurate and it takes a long time to train*

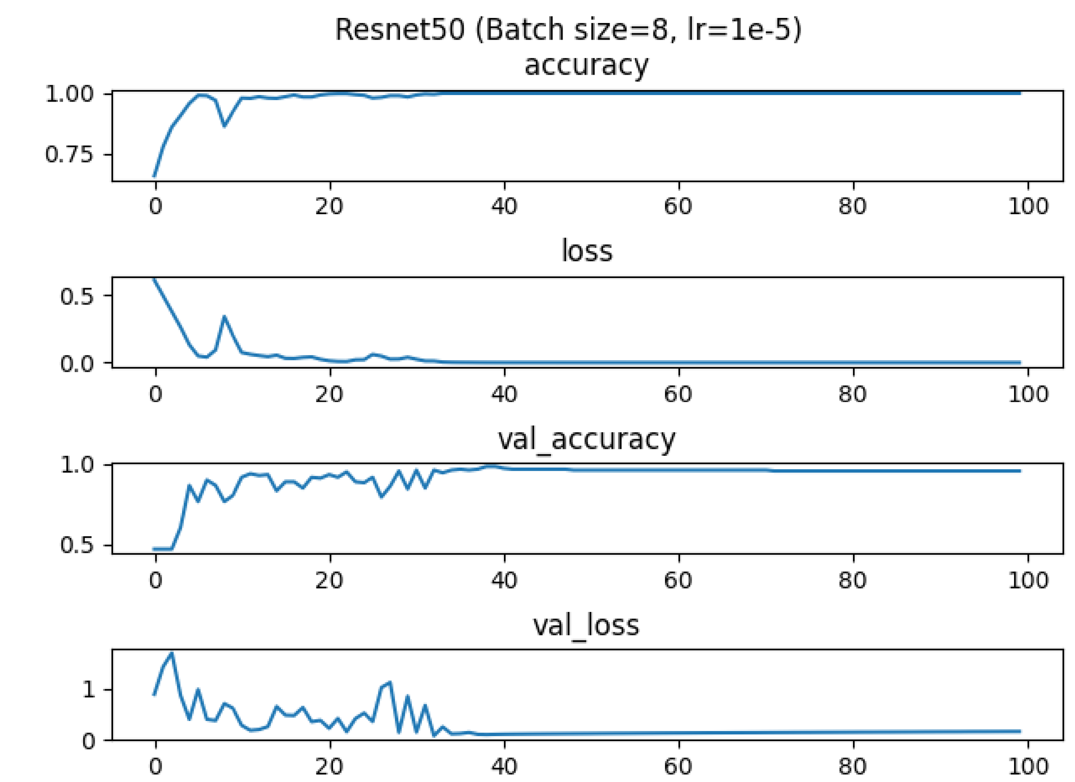


*Fig. 6: The results for these hyperparameters are much more promising*

### Result Validation

To ensure the model produced useful results, we applied it to fields that were part of neither the training nor validation data sets. The model successfully classified fields as waterlogged or non-waterlogged with a reasonable degree of accuracy.

An important edge case to test was fields that contain rivers and other waterways but are not waterlogged. Figure 4 is an example of such a field, and the model correctly classified it as not waterlogged. This indicates that the model is looking for a specific shape characteristic to waterlogged fields and not just any water appearing in the input image.

## FUTURE DIRECTIONS

### Waterlogging Susceptibility

Results from this project, applied to a large dataset, could be used to determine which geographic areas and climate conditions are most likely to result in waterlogged fields. These results would be useful to farmers, who could implement strategies to mitigate the risk of waterlogging. This could also be useful to researchers exploring the consequences of climate change.

### Impact of Waterlogging

As shown in figure 7, waterlogging can have a lasting impact on crop health. Combining crop yield data with waterlogging data gathered on a large scale through satellite imagery could provide further insights what impact waterlogging has on crops.

*Fig. 7: Even after the water recedes, waterlogging can cause plant disease[3]*

## REFERENCES

1. Adesina, Olufemi. (2018). Effect of Waterlogging on Soil. 10.13140/RG.2.2.25533.31202.

2. Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun. (2015). Deep Residual Learning for Image Recognition.

3. Staton, M. (2017, June 23). Assessing water damage to emerged soybeans. MSU Extension. https://www.canr.msu.edu/news/assessing_water_damage_to_emerged_soybeans.

## ACKNOWLEDGEMENTS

The authors thank the National Center for Supercomputing Applications for providing computing resources, the United States Department of Agriculture for providing Crop Data Layer information, and Planet Labs for providing satellite images.

ILLINOIS