Keep Summer Safe!

(Through Plant Disease Identification!)

Can an image recognition neural network be developed that looks at photos of crops and identifies whether or not they have a disease/sickness?

Further, if this works, could the neural network then be trained to specify what is likely wrong with said crop?

Short Summary

- Through the usage of convolutional neural networks I was able to create a high accuracy model that can look at images of rice, tomato, potato, and bell-pepper plants and identify whether they are healthy or not.
- I was unable to achieve my secondary goal and create a high accuracy multi-class model. I believe that it is possible to do this, but my own technological limitations hindered me.

Background Information

- The way crops are grown in most agricultural systems around the world means that if one plant ever gets sick on a farm plot, there is a significant chance that the rest of the plants of the same type in that plot will also get infected. This can be incredibly disastrous for farmers, who are often depending on stable seasonal yields to stay afloat.

- In places like the USA, where industrial monocrop systems of agriculture dominate food production, it has become extremely important to catch disease/sickness early and deal with it quickly.

Background Information 2

- Agriculture as a whole is already benefiting from, and could continue to benefit from the implementation of AI and automation. Harvesting and planting on a large scale can already be automated, but what about maintenance and ensuring healthy crop yields?

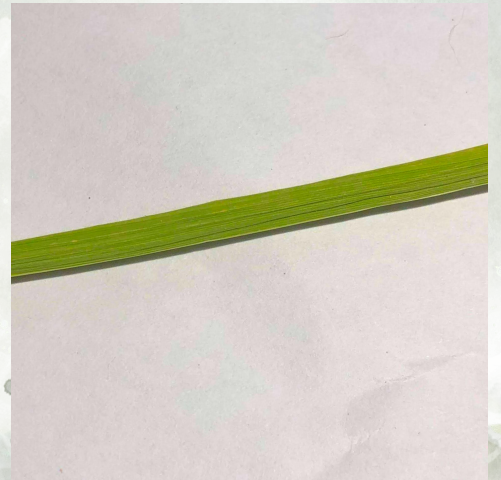
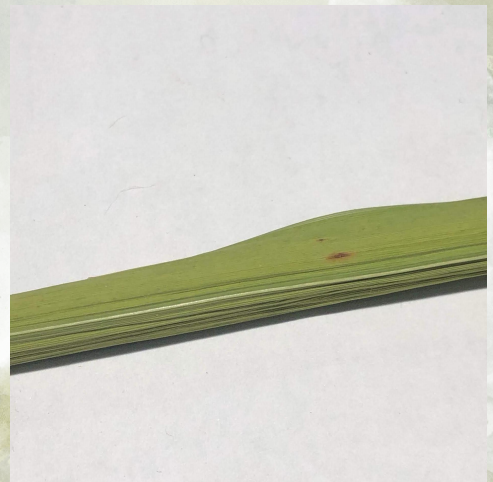
- The threat posed by climate change necessitates a strong toolkit to deal with the threat to agriculture posed by changes in our native ecosystems. A small increase or decrease in average temperature can have a huge impact on the spread of disease and sickness depending on the farming region on resides in.

Data Sources

- **Plant Village**
  - Please check them out! They have an amazing mission.
  - https://plantvillage.psu.edu/
- **Kaggle**
  - Specifically, this publicly available rice image datase https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset/kernels
- 23, 995 Images in Total with 19 different categories.





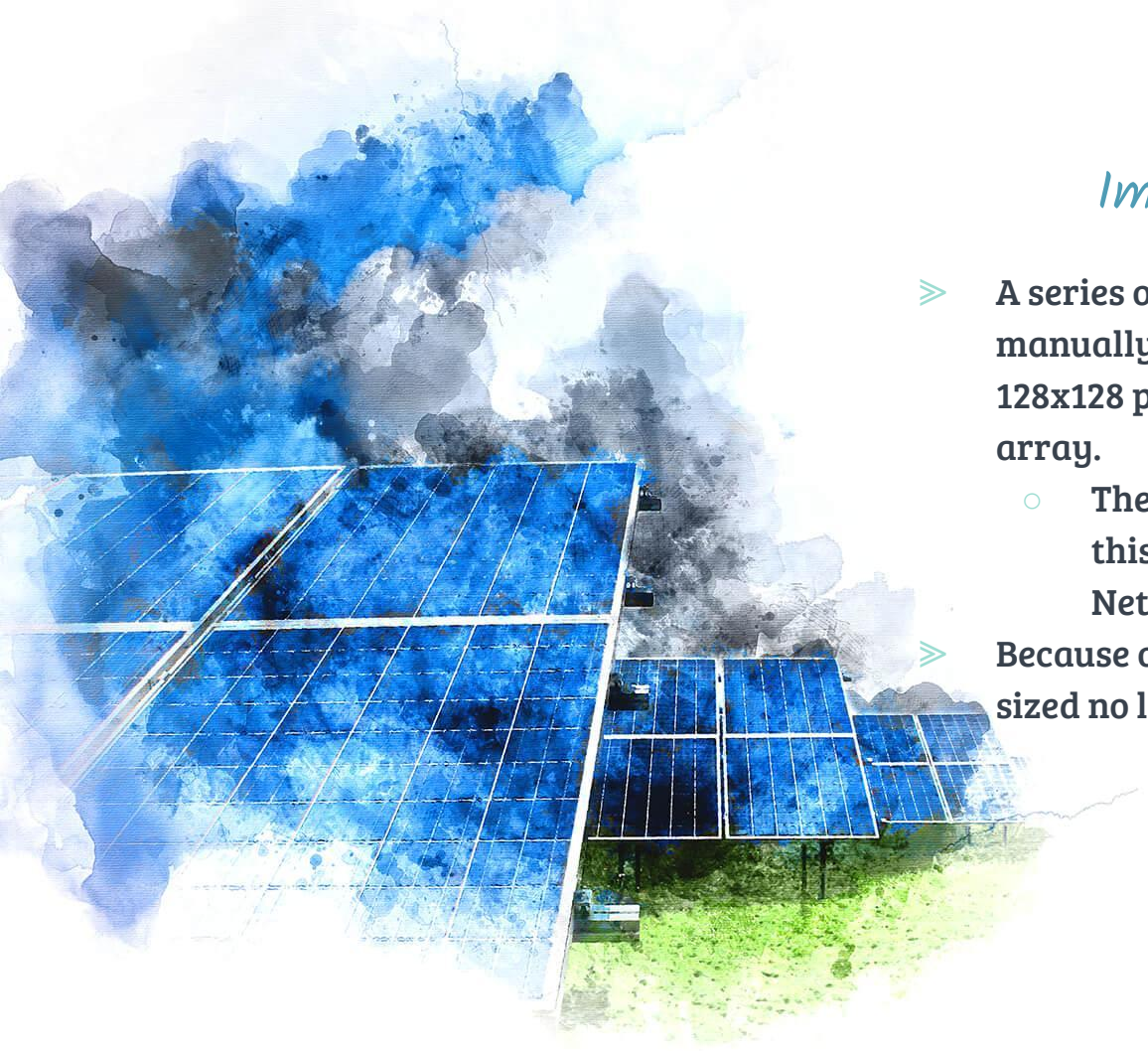


Image Pre-Processing

- A series of python functions were written to manually input each unique image, resize it to 128x128 pixels, and then convert it to a numerical array.
  - The image was converted to an array because this is what allows a Convolutional Neural Network to correctly process and learn.
- Because of computer limitations, images could be sized no larger than 128x128.

Two Different Approaches to making the CNN

Multiclass

- Using all 19 categories from the data to train the model to identify specifically for each category

Binary

- Labelling plant leaves as “healthy” and “not healthy” and training the model to identify whether or not an image is “healthy” or “not healthy”

Binary CNN

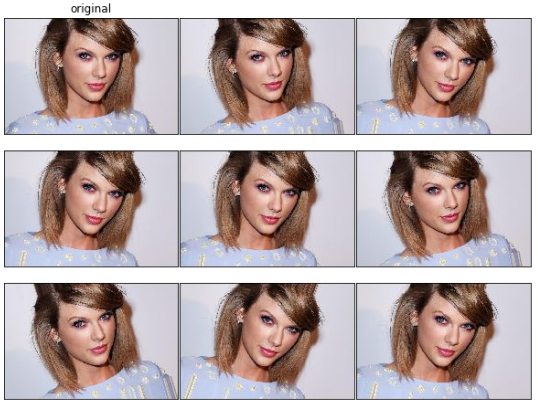
- As previously stated, the binary CNN I designed will simply state whether a plant image is “healthy” or “not healthy”.
- Best performing model with about 94% accuracy correctly predicting test images.
- Only able to reliably identify whether rice, potatoes, bell-peppers, or tomatoes are healthy or not healthy.
  - Needs to be trained on more different types of crop images.

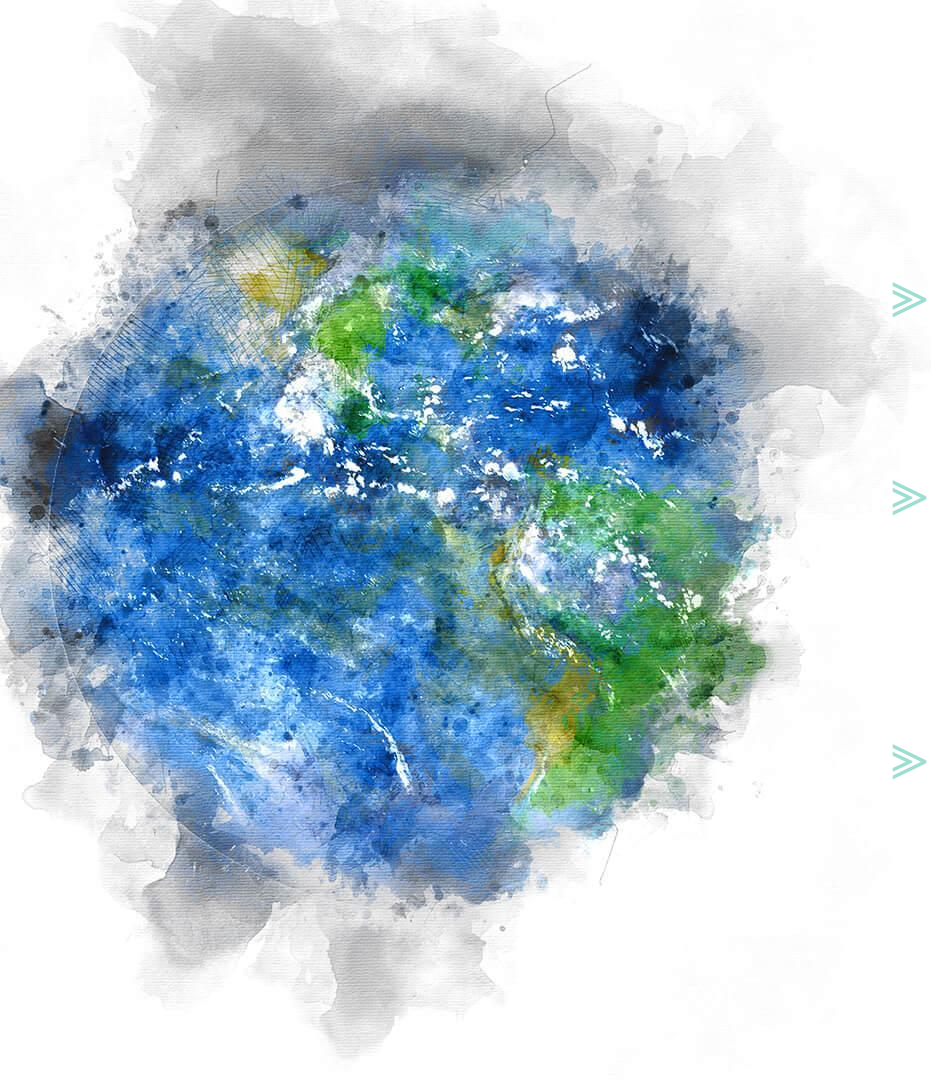
Multiclass CNN

- The goal with the Multi-Class CNN was to hopefully train the model to identify not only if a plant was healthy or sick, but also to say what type of plant it was and what sickness it had.
- Results for this model were mixed.
  - Training a CNN on just the “Plantvillage” dataset yielded an accuracy rate of about 89%
  - Training a CNN on just the “Rice” dataset yielded an accuracy rate of about 49%
- These scores could potentially be improved over time with more data as well as better computing capability.
- This model would be much harder to adapt to a large amount of plants.

Image-Data-Generator and CPU Problems

- One potential solution for the lack of images problems would be to implement and use a image-data-generator to manipulate my images and create a large amount of new data for the CNN to learn from.
- Sadly, this process requires a large amount of computer memory and my computer did not have enough memory to implement the task.
- In the absence of my own images, I'll let Tswift illustrate what this could do to my plant images.





Project Problems

- Computational Problems
  - Image Resizing
  - Image Generation
- Needs a lot more data to effectively solve initial project goal of generalizing “healthy” or “not healthy” to all types of crops.
- Effectively implementing a CNN on a desktop is very time consuming and getting the tools installed took up a lot more time than expected.

Problem-Solutions and Future Recommendations

Me adding layers to my Neural Network

- Neural Networks, especially CNNs and anything having to do with computer vision, are very complicated. In the future I have a lot more to learn about implementing them
  - For those technically inclined, I need to better understand pooling and dropout layers better as well as batch size.
- Learn AWS or some other cloud based software that can provide me much more CPU and GPU power.
- Gather more and more plant images!