

Article

Not peer-reviewed version

A Machine Learning Tool to Non-Invasively Detect Drought Stress in Plants

[Prisha Bhat](#) *

Posted Date: 22 October 2024

doi: 10.20944/preprints202410.1689.v1

Keywords: drought stress; rice; machine learning tool



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

A Machine Learning Tool to Non-Invasively Detect Drought Stress in Plants

Prisha Bhat

Plano East Senior High School, Plano, Texas, United States

Abstract: Drought stress is a major issue in agriculture, as it is becoming more prevalent due to climate change and can reduce crop yields significantly. Low water amounts, particularly at the flowering stage, limit plants from undergoing typical growth and reproductive development, including gametogenesis, fertilization, and embryogenesis. Therefore, it is important that rice growers have a method of quickly determining whether or not their crops are undergoing drought stress to enact mitigation strategies. This study aims to use machine learning computer vision to accurately detect drought stress in plants. Before I developed models, I conducted data pre-processing on images of *Setaria* plants with varying water contents. Then, I built several machine learning models, including K-Nearest-Neighbors, Decision Tree, Logistic Regression, and ResNet 18. I then assessed the effectiveness of these models by calculating their accuracy scores. I found that K-Nearest-Neighbors, Decision Tree, and Logistic Regression models obtained 80.2%, 70.7%, and 75.9% accuracy scores, respectively, while ResNet18 obtained a very high accuracy of 97.3%. Therefore, these models can serve as a promising future solution for non-invasively detecting drought stress in rice plants, potentially saving crop producers millions of dollars in yield.

Keywords: drought stress; rice; machine learning tool

Introduction

Over 50% of the world's population is dependent on rice crops. However, drought stress threatens to decrease rice yields by up to 92%, depending on severity (Ahmad et al., 2022). Drought stress also increases the susceptibility of rice plants to arsenic stress. Generally, periods of drought are accompanied by heat stress, and this heat breaks bonds between arsenic and other minerals, releasing more free arsenic into the soil (Muehe et al., 2019). Current strategies for drought detection are generally very costly, labor-intensive, and not scalable. Therefore, there is currently a need for a non-invasive method to accurately determine if plants are drought-stressed or not. Machine learning (ML) using computer vision therefore serves as a promising solution to detect drought stress in rice plants without needing to conduct invasive, expensive, and time-consuming wet-bench assays.

This study employed several different baseline and advanced ML models to determine the most effective one, including Logistic Regression, K-Nearest-Neighbors, Decision Tree, and ResNet 18.

Background

Plant Physiological Responses to Drought Stress

Plants respond to drought through a series of physiological changes aimed at conserving water. One of the primary responses is stomatal closure, which reduces water loss through transpiration however also limits carbon dioxide uptake and photosynthesis. Drought stress triggers the accumulation of abscisic acid (ABA), a hormone that plays a key role in stomatal regulation and other drought responses (Bharath & Gahir, 2021). Additionally, osmotic adjustment occurs as plants accumulate solutes like proline and sugars, helping to maintain cell turgor and protect cellular structures. Over time, prolonged drought can lead to oxidative stress, causing damage to proteins, lipids, and DNA, and potentially resulting in cell death if the stress is severe and prolonged. These

all result in observable changes. One of the first visible signs of drought stress is wilting, as reduced water availability leads to less turgor pressure in plant cells, thus causing plant leaves and stems to droop. Another common indicator of drought stress is leaf discoloration, as leaves turn yellow or brown due to the degradation of the green pigment chlorophyll within cells (Seleiman et al., 2021). In very severe cases of drought, entire sections of plants will die, starting with the tips of leaves and branches.

Current Strategies for Drought Detection

Current techniques for detecting drought stress in plants involve both direct and indirect methods. Direct methods include measuring leaf water potential using a pressure chamber or psychrometer, which measures the water content of the plant (Rockwell et al., 2014). Stomatal conductance can be assessed with porometers to gauge the extent of stomatal closure (Jones, H. G. 2006). Indirect methods involve remote sensing technologies, such as thermal imaging and multispectral or hyperspectral imaging, which detect changes in leaf temperature and reflectance associated with drought stress (Zarco-Tejada, et al., 2002). Soil moisture sensors can also be used to measure water availability around the plant's root zone. However, limitations of these techniques are that they are often labor-intensive, time-consuming, and not feasible for large scale monitoring (Turner, N. C., & Long, M. J., 1980). Particularly, remote sensing may be useful for large plots of land, but is not very precise when examining individual plants. Soil moisture sensors also do not directly measure plant stress.

Current Machine Learning Strategies for Drought Detection

While machine learning tools for detecting drought stress hold promise in agriculture, they face several significant limitations. One major challenge is their reliance on high-quality, diverse datasets, which are often unavailable in many regions, particularly in developing countries. This data scarcity can lead to inaccurate or unreliable predictions. Furthermore, these tools often struggle with generalization across different geographic areas and crop types, resulting in varying accuracy and effectiveness. The complexity of the algorithms also poses a barrier, as farmers may require substantial training to effectively integrate these tools into their existing practices. Additionally, the high computational resources needed for these models can be a limitation, especially for small-scale farmers with limited access to advanced technology. These factors collectively limit the widespread adoption and effectiveness of machine learning in drought stress detection.

Dataset

A public-access dataset containing 1,156 visible close-up images of *Oryza sativa* (Rice) leaves at various stages of wilting due to nutrient deficiency was obtained from Kaggle. A workable dataset of *O. sativa* images wilting due to drought stress alone was not available, and since literature suggests the correlation between symptoms of nutrient deficiency and drought stress in rice, this dataset was used as an alternative. The original dimensions of the images were approximately 3834 x 310 pixels. However, I resized the images during data pre-processing.

Using the `train_test_split` function from `scikit-learn`, I split the dataset: 50% of the images were used for training, 30% of the images were used for validation, and the last 20% were used for testing. Before images were classified, I created NumPy arrays 'X' and 'y' for the images and their nutrient amounts, respectively.

Methodology/Models

Four machine learning models were used to classify the plant images: Logistic Regression, Decision Tree, K-Nearest-Neighbors, and ResNet18.

Classification Using Logistic Regression

Logistic Regression (LogReg) machine learning models sorts data by calculating the likelihood of a datapoint belonging to a certain class. Though not very advanced, LogReg models are space-efficient and good at classifying linearly seperable data.

Classification Using Decision Tree

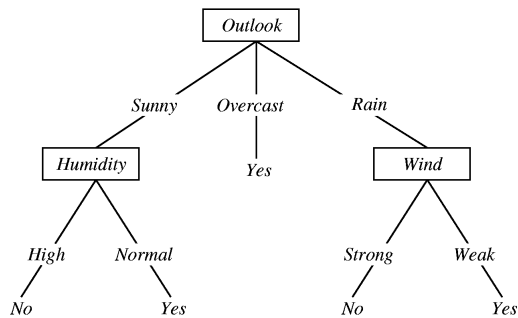


Figure 1. Diagram of Decision Tree.

Decision Tree models contain various layers that the images are passed through. Decision Trees have a hierarchical structure made up of nodes and branches to model every possible decision and outcome produced. The nodes act as decision makers inside the model that actually separate and classify the image data. Ideally, there should be as few nodes as possible in a Decision Tree model to avoid overfitting the model to the testing data.

Classification Using K-Nearest-Neighbors

K-Nearest-Neighbors (KNN) is a model used for classification tasks. It classifies data points based on the majority class of their closest neighbors, and this “closeness” is determined by calculating the distance between a target point and all other points in a dataset. The “K” in KNN is a variable representing the number of nearest neighbors considered.

Classification Using RESNET18

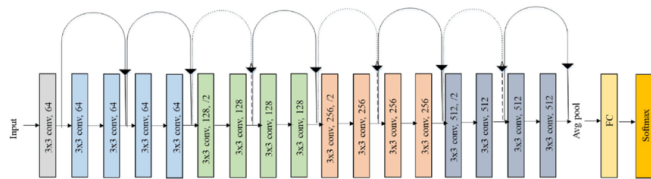


Figure 2. Diagram of ResNet18 Model Architecture.

ResNet18 is a deep learning model used for image classification tasks, and does not flatten image data, making it optimal for use to detect drought stress from images of plants. ResNet18 consists of 18 layers, including convolutional, pooling, and fully connected layers. The model uses residual connections that allow it to bypass certain layers, mitigating the “vanishing gradient” problem common in deep networks. The vanishing gradient problem occurs when gradients used to update the network become very small as they are back-propagated from output layers to earlier layers, making it difficult for the network to effectively update weights, which slows the training process. Despite being less complex than other models within the ResNet family, such as ResNet50. ResNet18 can achieve high accuracy and balances speed and performance well.

Results and Discussion

Drought stress in crops is a major issue that can severely threaten the yield and quality of crops, and thus the lives of billions of people. Therefore, it is important for crop-growers to be able to detect drought stress in plants early on to employ effective mitigation strategies.

In this study, I found that baseline models were somewhat effective while the ResNet18 model was very effective. The Logistic Regression model obtained an accuracy of 75.9%. The K-Nearest-Neighbors model obtained an accuracy of 80.2%. The Random Forest model obtained an accuracy of 82.8%. The Decision Tree model obtained an accuracy of 70.7%. ResNet18 obtained an accuracy of 97.3%, which is very high especially compared to existing models.

In the future, this tool can be implemented to help crop-growers in drought-prone areas monitor their crops to ensure that their health is not degrading. The tool can also be used when testing wet-bench treatments to improve drought tolerance in plants.

For further research, I can develop more advanced models and test the models on different plant species and genera from *Setaria* that was used in this model. For actual implementation, in the future, this tool can be used to help crop-growers in drought-prone areas monitor their crops to ensure that their health is not degrading. The tool can also be used when testing wet-bench treatments to improve drought tolerance in plants.

Acknowledgments: I would like to thank Juyon Lee for providing me guidance throughout the process. The dataset for this study was obtained from Kaggle.

References

1. Ahmad, H., Zafar, S. A., & Naeem, M. K. (2022, March 23). Impact of Pre-Anthesis Drought Stress on Physiology, Yield-Related Traits, and Drought-Responsive Genes in Green Super Rice. *Frontiers in Genetics*, 13(2022). 10.3389
2. Bharath, P., & Gahir, S. (2021, March 4). Absciscic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress. *Frontiers in Plant Science*. PubMed. 10.3389
3. Jones, H. G. (2006, September 15). Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. *J Exp Bot*. PubMed. 10.1093/jxb/erl118
4. Muehe, E. M., Wang, T., & Kerl, C. F. (2019, November 1). Rice production threatened by coupled stresses of climate and soil arsenic. *Nature Communications*. <https://doi.org/10.1038/s41467-019-12946-4>
5. Rockwell, F. E., Holbrook, N. M., & Strock, A. D. (2014, February 26). The competition between liquid and vapor transport in transpiring leaves. *Plant Physiology*. PubMed. 10.1104/pp.114.236323
6. Seleiman, M., Al-Suhaibani, N., & Ali, N. (2021, January 28). Drought Stress Impacts on Plants and Different Approaches to Alleviate Its Adverse Effects. *Plants (Basel)*. PubMed. 10.3390/plants10020259
7. Turner, N. C. (1980, January). Errors Arising From Rapid Water Loss in the Measurement of Leaf Water Potential by the Pressure Chamber Technique. *Australian Journal of Plant Physiology*. ResearchGate. 10.1071/PP9800527
8. Zarco-Tejada, P. J., Miller, J. R., & Mohammed, G. H. (2002, September). Vegetation stress detection through chlorophyll a + b estimation and fluorescence effects on hyperspectral imagery. *Journal of Environmental Quality*. PubMed. 10.2134/jeq2002.1433

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.