PROJECT 1 REPORT PAPER

Joseph Martínez

# 1. Problem Definition

The Plant Pathology 2021 - FGVC8 Kaggle competition [1] aims to address the pressing need for automated identification and classification of foliar diseases in apple trees. Currently, disease diagnosis in apple orchards relies heavily on manual scouting by humans, which is not only time-consuming but also financially burdensome.

Despite the potential of computer vision-based models in disease identification, there exist significant challenges that must be overcome. Variations in visual symptoms of diseases across different apple cultivars, as well as environmental factors such as leaf color, morphology, and lighting conditions during image capture, pose substantial obstacles for accurate disease classification.

Building upon the success and lessons learned from the Plant Pathology 2020-FGVC7 challenge, 2021's competition presents a more extensive dataset consisting of approximately 23,000 high-quality RGB images of apple foliar diseases. The dataset includes diverse backgrounds, leaf maturity stages, and varying focal camera settings, providing a more comprehensive representation of real-world scenarios.

The primary objective of this competition is to develop machine learning-based models capable of accurately classifying leaf images into specific disease categories and identifying individual diseases amidst multiple symptoms present on a single leaf image.

By leveraging the provided dataset and advancing state-of-the-art machine learning techniques, participants are tasked with creating models that can significantly enhance the efficiency and effectiveness of disease diagnosis in apple orchards. Through this competition, we aim to foster innovation in agricultural technology and contribute to the sustainable management of plant diseases.

# 2. Methods

Provide detailed description on your methods, machine learning model/architecture, data processing/preprocessing, feature selection/extraction/engineering, implementation details (pseudo code, flow chart, etc.), and training/testing/validation process. If you build your program on top of somebody else’s program, make sure you reference. It is important to specify clearly what you have implemented/contributed.

# 3. Results

Describe the performance of your ML model, what is the accuracy, precision, recall, etc. for task 1. Analyze the results you get, for example, how do methods/strategies lead to a better solution? What do you learn from this project? Use figures, charts, and tables to assist your analysis. You can also compare your results with the Kaggle leaders.

# 4. Conclusion

Provide your conclusion. Do you achieve your goal? Why or why not? Is there a way to do it better?

# 5. References

1. Thapa, *Plant Pathology 2021 - FGVC8*. Kaggle.