A classic approach to detect potato diseases

Jeffri Erwin Murrugarra Llerena
National University of Trujillo, Peru

Latinx in AI

Introduction

- We examine how to detect pests (**Phytophthora infestans**, **Alternaria Solaris**) in potato leaves.
- We used a classic approach combining image processing and machine learning algorithms.
- We obtain a 96.3% precision accuracy for the best classifier.

Motivation

- Peru has 3000 varieties of potatoes and constitutes **25%** Gross domestic product (**GDP**) and it is the main resource of nutrition in Andean cities.
- The loss of crops of potatoes by pests is **20%** to **40%** per year according to the United Nations Food and Agriculture Organization (**FAO**), and in Peru is **15%** to **46%.** In order to tackle this issue, we develop a tool to early detect pests in potato crops.

Key idea

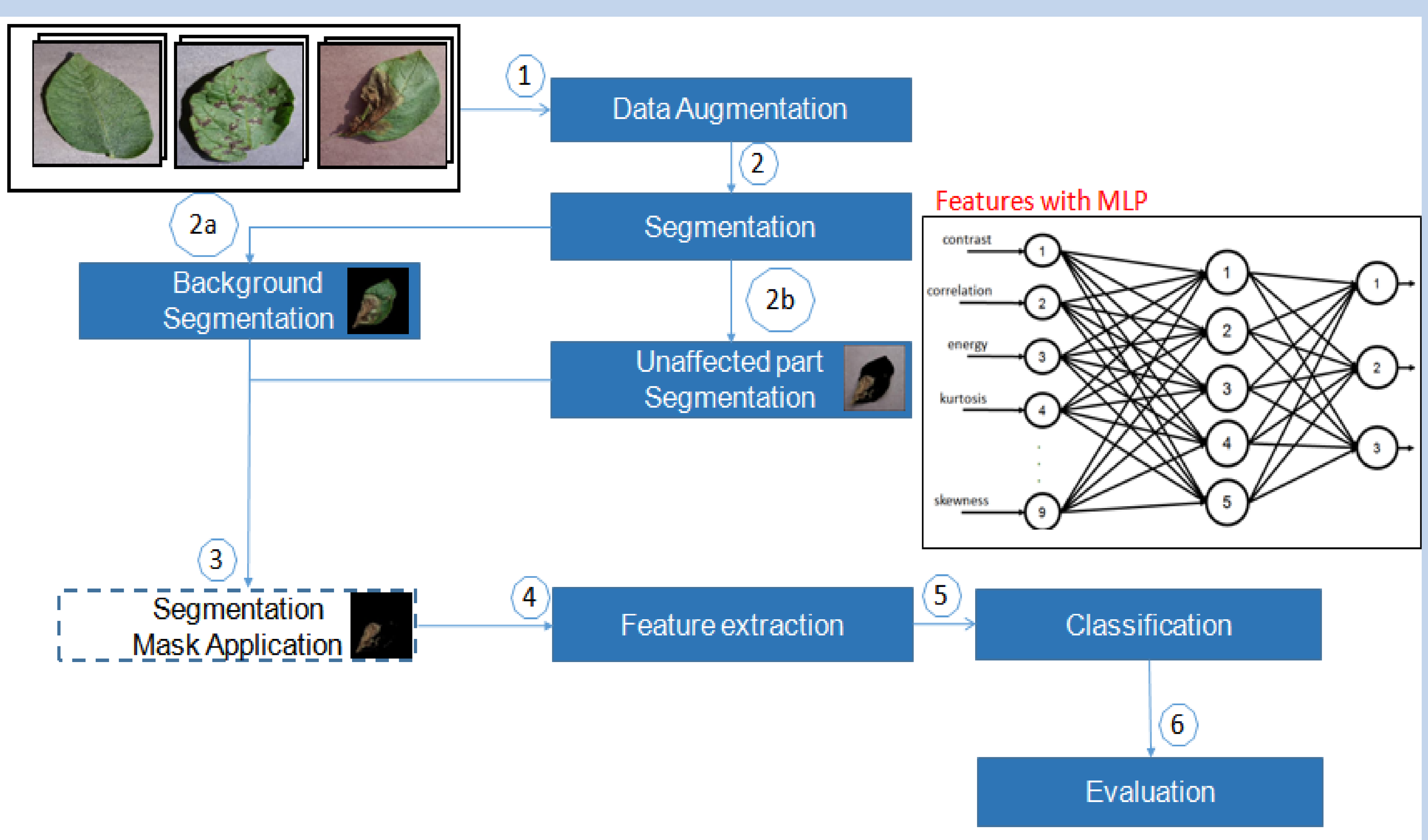
- Automatically detect pests in potato from a leave image.

Classes
•Good conditions
•Alternaria solaris
•Phytophthora infestans

Related work

- Prior work detects diseases in olives **[4]** or apples **[5],** and similar work evaluates if a fruit is ripe **[6]** using a fruit image. Instead in our case, we use a leave image.

Approach



- We increase our dataset size using the following transformations: rotations, turns, and reflections.
- In Background segmentation, we work in **HSV** space in the channel S because greater grayish hue is easily differentiable and that is present in our image's background, and we apply Otsu **[1]** algorithm. Finally, we apply **morphological operations** (dilation and erosion) to soften the image.
- Then, we remove **holes in the image** applying dilation with a matrix of zeros with a padding of 1.
- In the unaffected part segmentation: we work in **l*a*b*** because is easy to extract the green color (a*< 0).
- We extract 9 features: **contrast**, **correlation**, **energy**, **homogeneity**, **mean**, **standard deviation**, **entropy, kurtosis** and **skewness.** The first four features were extracted from the co-occurrence matrix and the others from the histogram.

Experimental setup and Dataset

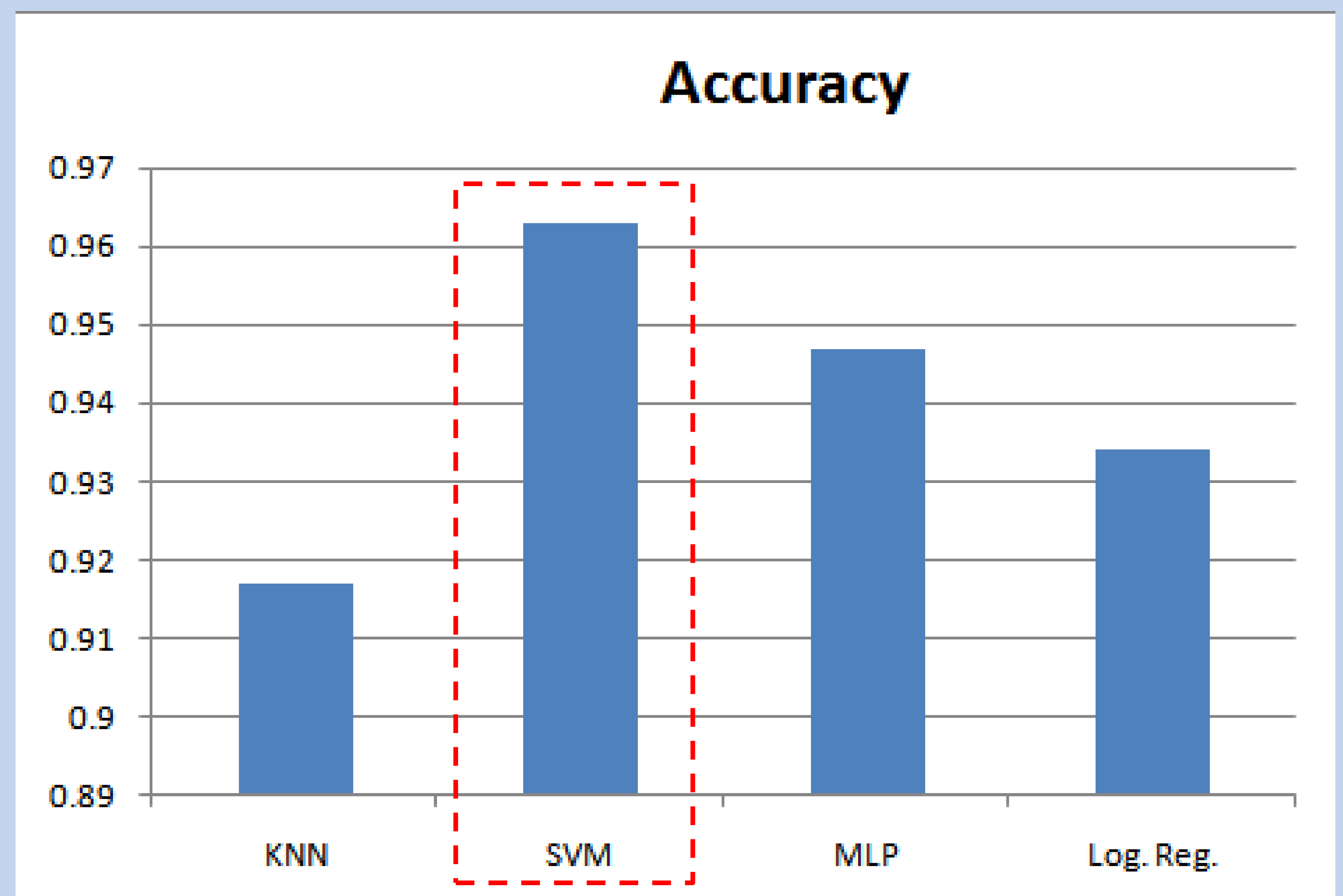
We use Plant Village Dataset **[3]**

We split our data in:
- 60% for training.
- 20% for validation.
- 20% for test.

152 Good conditions leaves
152 Alternaria Solaris leaves
152 Phytophthora infestans leaves

Evaluation

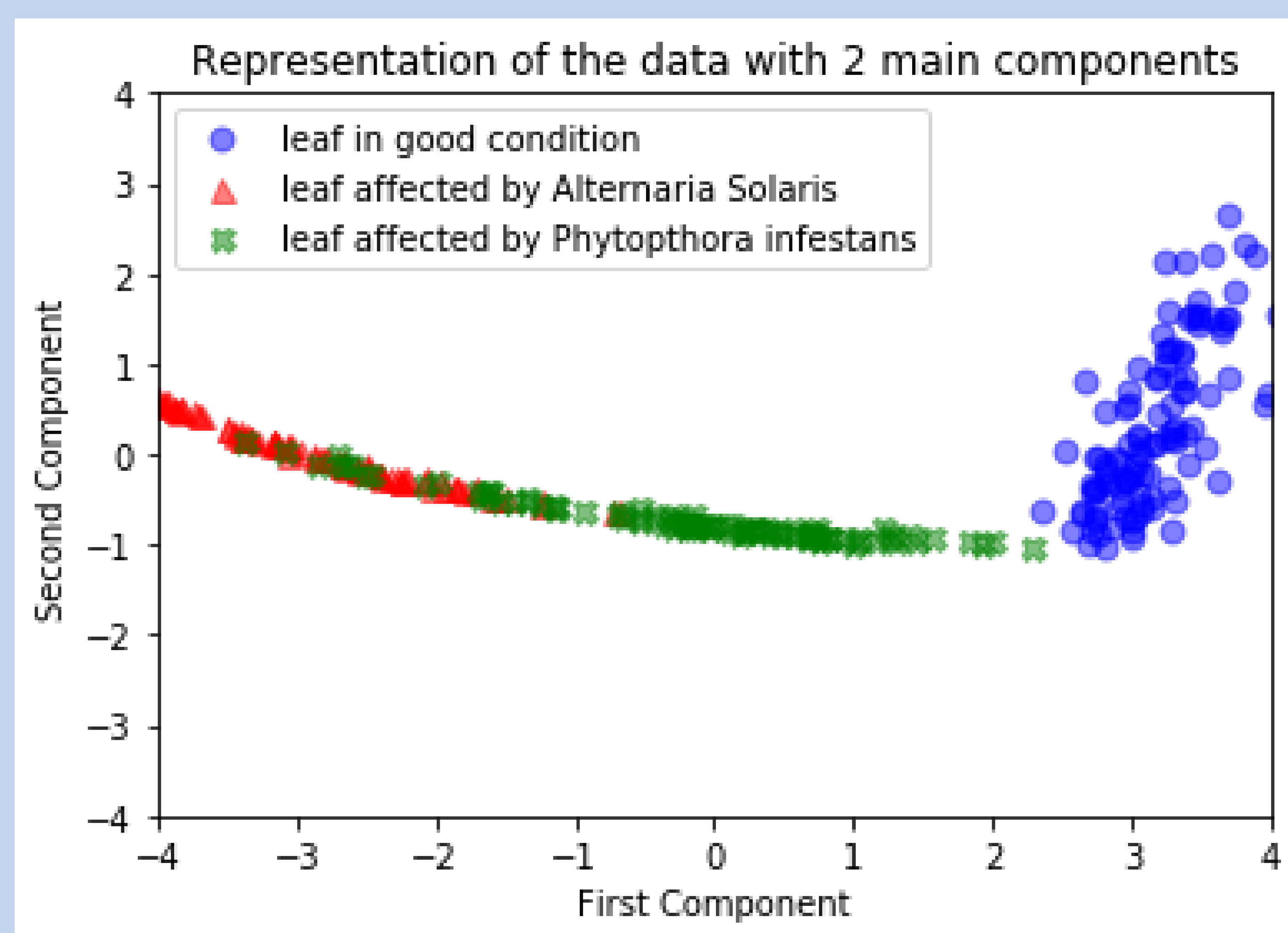
We compare four models, the best parameters were selected with the validation set:
- Five Nearest Neighbors using **Euclid Distance**.
- Support Vectors Machine with a **rbf** kernel, one – rest classifier, penalty 1, gamma 1/ number of examples.
- MLP with 9 neurons in the input layer, 5 neurons in the hidden layer, 3 neurons in the output layer using **backpropagation** with **Adam optimizer [2]**, **ReLU** activation function without penalty and randomized weights.
- Logistic Regression with **L2** regularization and **LBFGS** optimizer.



- We compare these methods with accuracy because the data is balanced, and SVM has the best performance.

Qualitative results

- We want to analyze and interpret better our performance. Hence we visualize our features using Principal Component Analysis (PCA).



- We find that two main components obtain **96%** of representativeness of the data.
- Also, we observe that Class 1 (Potato in good condition) is **linearly separable** from the other 2 classes, also, the other two classes present low overlap.
- Considering these two facts, we observed that our classes are easy to separate, which ensures **high performance**.

Contributions

- We provide a classic **approach for detecting potato plants** affected by the Phytophthora infestans and Alternaria Solaris on the Plant Village dataset **[3]**.
- We expect to apply this work on many communities in Latin America and achieve a high impact on agriculture.

Future work

- We plan to use **deep learning** to find better feature representation and compare with traditional methods, deep learning can provide even better performance.
- We plan to **expand this algorithm** for a wider variety of pests and other plants.

References

[1] Nobuyuki Otsu (1979). A threshold selection method from gray-level histograms. In IEEE Transactions on Systems
[2] Kingma Diederik & Adam Jimmy (2015). A method for stochastic optimization. International Conference on Learning Representations
[3] Hughes David & Salathé Marcel (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics.
[4] Carranza Freddy & Murrugarra Nils (2007). Detection of fisheye disease using image processing. Peruvian Conference on Computing
[5] Woodford Brendon (1999). Fruit analysis using wavelets
[6] Murrugarra Joseph (2012). Classification of the state of maturity of apples by artificial vision. Compuscientia