Plant Disease Classification

Fundamentals of Data Science

Alessio Iacono

Leonardo Rocci

Mohammadreza Heibati

Antonio Pagnotta

github.com/cycileo/FDS-Project

Sapienza University of Rome

TASK AND MOTIVATION

Task Statement:

- Develop a machine learning model to classify plant diseases based on leaf.

Motivation:

- Early and accurate detection of plant diseases prevents crop losses and ensures food security.
- Farmers rely on manual inspection, which is time-consuming and error-prone

→

MODELS AND TOOLS

Dataset:

- The PlantVillage dataset contains 54,000 labeled images across 38 categories, which include both healthy and diseased plants.
- The dataset covers 14 plant species.
- The dataset is split into 80% for training, 10% for validation, and 10% for testing, with the split maintaining the original class proportions.

Tools and Techniques:

- ~~Data Augmentation: rotation and brightness variation to simulate variation in leaf positioning and lighting~~ (in future)
- Preprocessing: normalization to improve model convergence.
- Model: Custom-built CNN for plant disease classification, and fine-tuned version of ResNet.
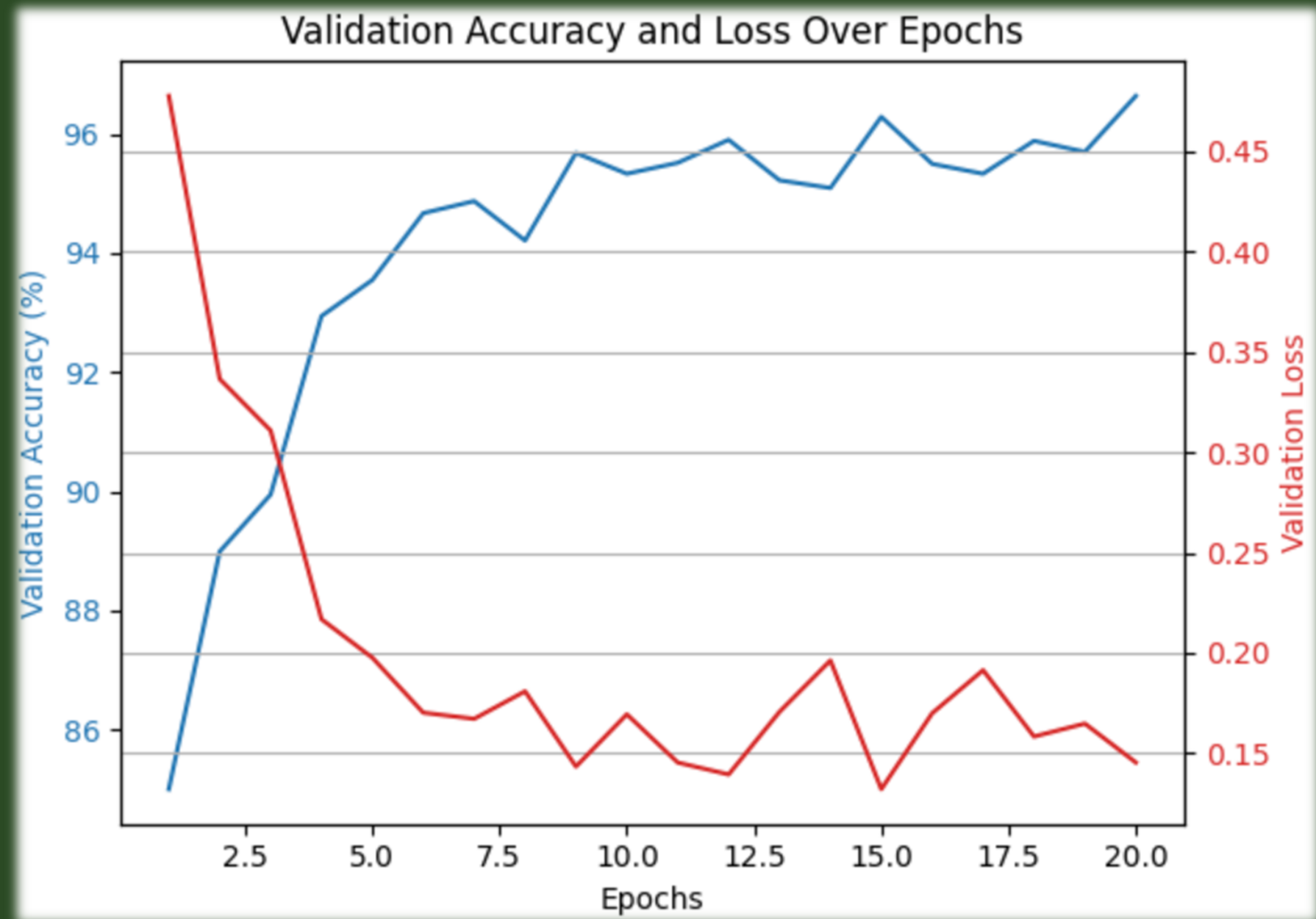
RESIZE OR NOT RESIZE?

We chose the lower image width as it provided similar accuracy while significantly reducing running time, making it the optimal choice.
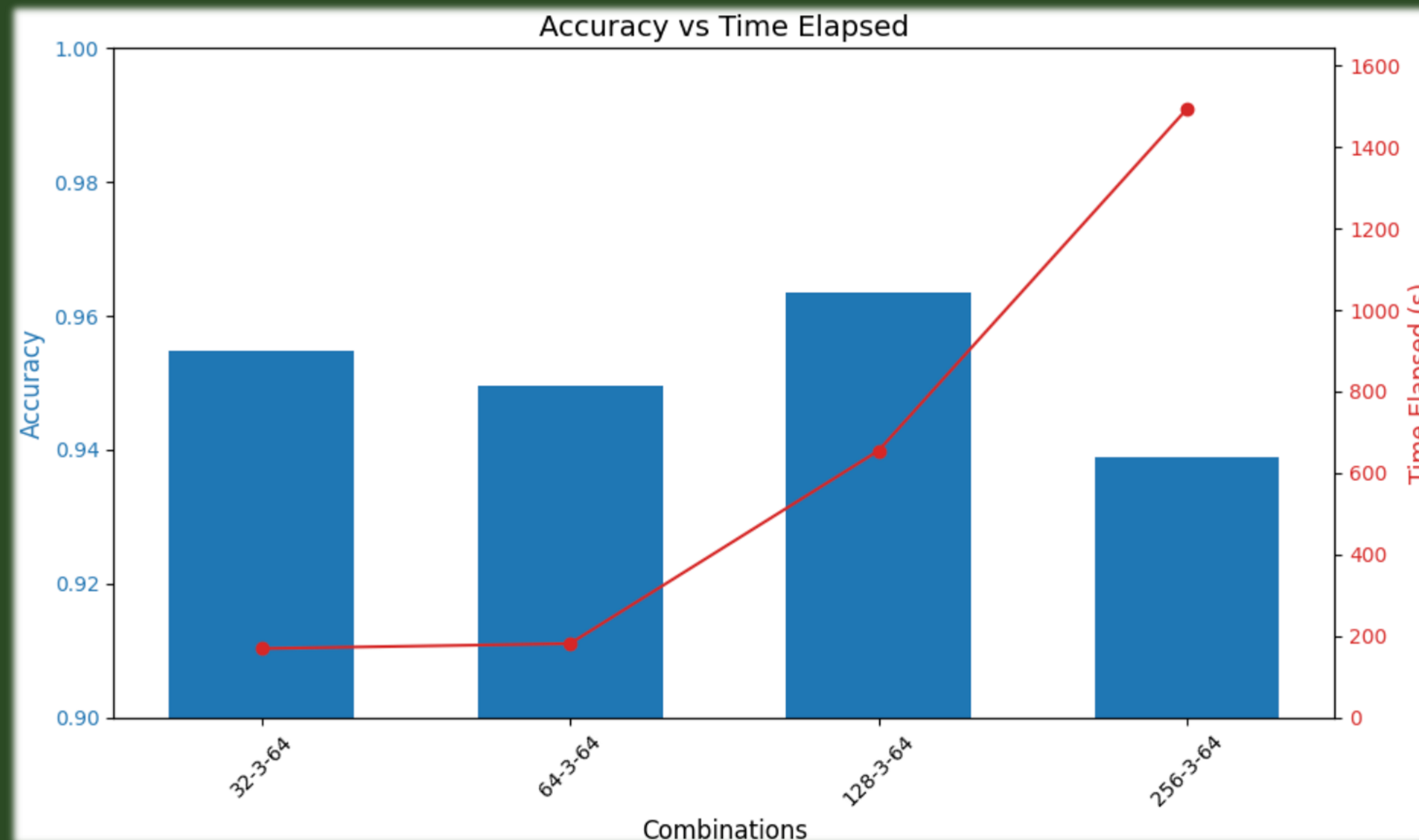
OUR MODEL

- IMAGES WIDTH: 32
- LAYERS: 3
- TRAINING TIME: 3 min.
- EPOCHS: 19

- EVALUATION:
  ACCURACY: 95%
  F1 SCORE: 0.95

Validation Accuracy and Loss Over Epochs

EVALUATION ON REAL-WORLD SCENARIOS

We curated a dataset of approximately 200 images sourced from Google to evaluate the model's performance in real-world scenarios.

PlantVillage

Google Images

EVALUATION ON REAL-WORLD SCENARIOS

Given the homogeneity of conditions in the training dataset, the performance was underwhelming.

Our Model

Accuracy: 11%
F1 Score: 0.09

Fine-tuned ResNet

Accuracy: 24%
F1 Score: 0.20

Future Goal: Aim to achieve better performance by incorporating data augmentation techniques to enhance the diversity of the training dataset and by utilizing segmented images to focus the model on relevant features.

THANK YOU FOR THE ATTENTION

Relevant Links

- Dataset:
  PlantVillage dataset

- Relevant Paper:
  ITM Conference: Plant Disease
  Detection and Classification
  using Deep Learning