Tomato Leaf Disease Classification

CSCI 5525 Final Project Fall 2023

Tony Zhang

Introduction

- Tomatoes are one of the most popular vegetables
- 256 million tons produced in 2021 [1]
- Cheap and healthy
- Effective large scale production can reduce malnutrition[2]

How The Tomato Transformed The European Diet - Epicure & Culture (epicureandculture.com)
Maharashtra: Tomato Farmer Profits Rs 2.4 Crore In Current Season Amid Price Rise (outlookindia.com)

Disease Detection and Treatment

- Difficult to manually search for diseases
- Large amounts of pesticide used in treatment
- Pollution of food and water systems [3]
- Many diseases show in the plant's leaf
- Automatically monitor crop health via computer vision

Save Your Tomato Plants From These 15 Common Diseases (thespruce.com)

Problem and Data Set

Machine Learning Problem

- Train model to detect diseases from image of tomato leaf
- Discoloration, shape, texture, etc.
- Wide variety of ways a disease can manifest

Data Set

- Tomato leaf disease detection (kaggle.com)
- 9 diseases + healthy = 10 classes
- 1000 training images per class
- 100 testing images per class
- Noise added

Algorithms Trained

CNN - TomaNet

- Trained with custom adam optimizer
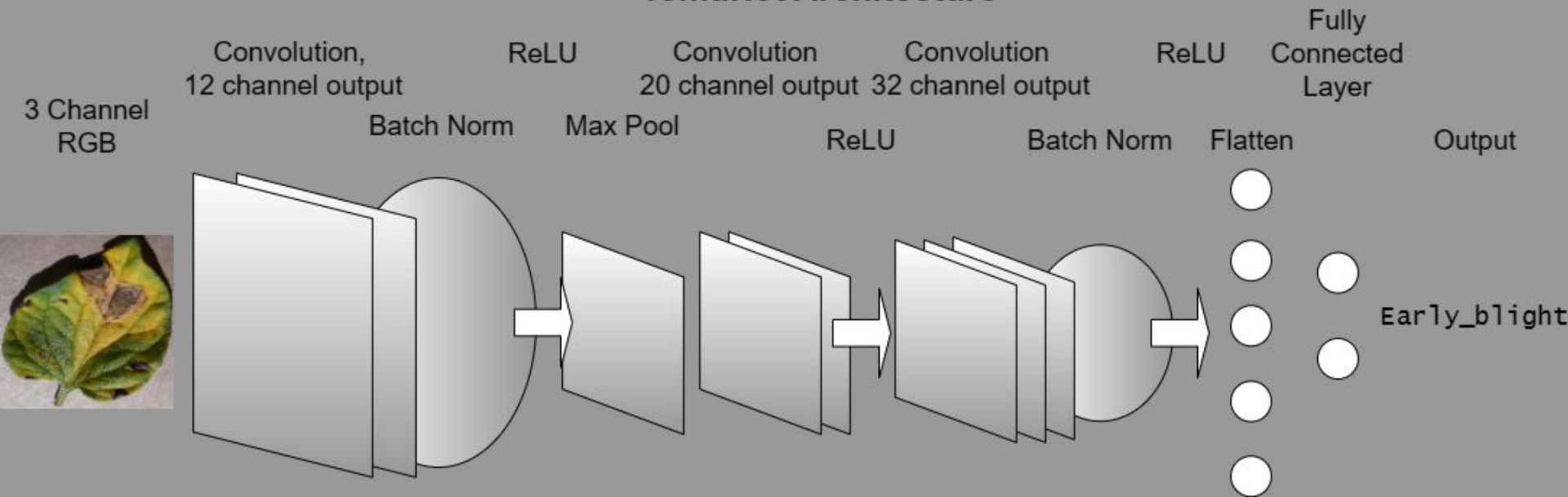- Also trained on PyTorch's default adam implementation

Random Forest

- 100 trees

Logistic Regression

- Trained with stochastic gradient descent

TomaNet Architecture

3 Channel RGB

Convolution, 12 channel output

Batch Norm

ReLU

Max Pool

Convolution 20 channel output

ReLU

Convolution 32 channel output

Batch Norm

ReLU

Flatten

Fully Connected Layer

Output

Early_blight

Results

Table 1: Test set accuracy for each disease by algorithm

Accuracy by Algorithm	TomaNet PyTorch Adam	TomaNet Custom Adam	Random Forest	Logistic Regression
Healthy	0.83	0.76	0.66	0.36
Bacterial Spot	0.95	0.89	0.83	0.59
Early Blight	0.81	0.31	0.40	0.52
Late Blight	0.81	0.63	0.54	0.46
Leaf Mold	0.74	0.58	0.55	0.58
Septoria Leaf Spot	0.81	0.73	0.57	0.67
Spider Mites	0.85	0.82	0.59	0.67
Target Spot	0.89	0.75	0.46	0.76
Mosiac Virus	0.97	1.0	0.69	0.78
Yellow Leaf Curl Virus	0.95	0.92	0.79	0.58
Overall	0.86	0.74	0.61	0.60

Results

Table 2: Training loss and test set accuracy by epoch for TomaNet trained with my custom implementation of Adam

	Training Loss	Test Set Accuracy
Epoch 1	57.431	0.731
Epoch 2	24.614	0.586
Epoch 3	4.892	0.739
Epoch 4	1.577	0.518
Epoch 5	1.425	0.546
Epoch 6	1.359	0.718
Epoch 7	1.201	0.693
Epoch 8	1.042	0.700
Epoch 9	1.114	0.692
Epoch 10	1.046	0.631

Live Demo

References

[1] Gleaned from data provided by FAOSTAT https://www.fao.org/faostat/en/#home

[2] Global Alliance for Improved Nutrition. Tomatoes, the world's most popular vegetable. https://www.gainhealth.org/sites/default/files/publications/ documents/advocacy-brief-tomatoes-the-worlds-most-popular-vegetable.pdf

[3] Doris Sande, Jeffrey Mullen, Michael Wetzstein, and Jack Houston. Environmental impacts from pesticide use: a case study of soil fumigation in florida tomato production. International journal of environmental research and public health, 8(12):4649–4661, 2011