

**Indian Institute of Science Education and Research
Mohali**

**Introduction to Data Science
Instructor: Vishal Bharadwaj**

Leaf Classification

Group 1

Amirr, ms20160
Sharat Sapna Anuj, ms20094
Sanjana Satish, ms20164

Due Date: November 7, 2023
Submitted: November 7, 2023

1 The Question

1. Take photos of the leaves from three different trees.
2. Create a model to classify them (including an unidentified tree).
3. Predict the type of leaf from a photo using the model.

2 Our Approach

To address this question, we employ a Convolutional Neural Network (CNN) model and apply it to a dataset comprising approximately 340 data points, along with augmented data points (10x), resulting in a total dataset of 3400 data points. We utilize CNNs due to their suitability for analyzing visual inputs, making them a scalable choice for image and object identification tasks.

3 Brief Explanation of the Code

Our model consists of seven convolution layers and three fully connected layers. All layers contain 64 neurons, except for the last convolutional layer, which has 30 neurons (corresponding to the 30 species in our dataset). After extensive experimentation, we determined the following hyperparameters:

1. Regularization - 0.01
2. Batch size - 6
3. Dropout probability - 0.01

We use ReLU and softmax activations to introduce non-linearity in our layers and prevent overfitting. Additionally, we employ ADAM optimization to automatically adjust the learning rate. We utilize sparse categorical cross-entropy as our loss function, which is well-suited to this problem.

Our model is trained for 100 epochs, and we split the dataset into an 80/20 ratio for training and validation.

The accuracy achieved by our model is 93.4

4 Important Links

- Leaf Data Set
- Notebook
- Git link

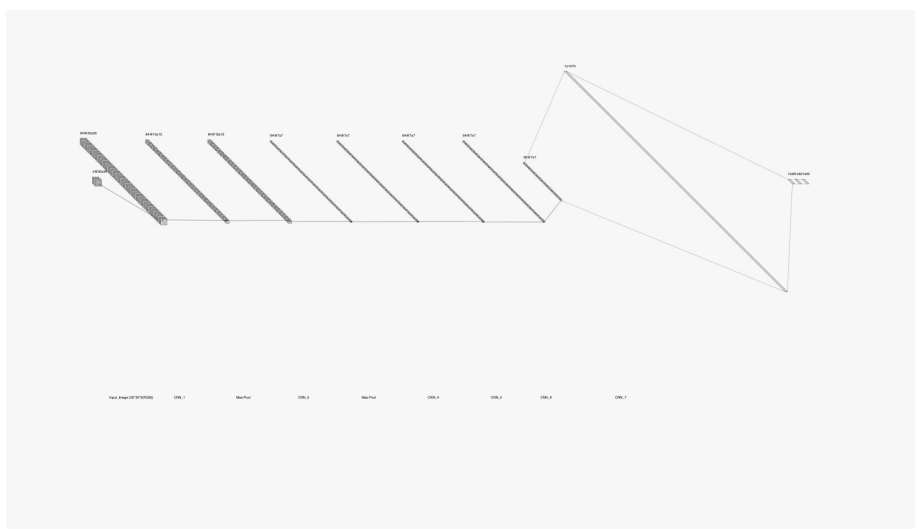


Figure 1: Schematic of the CNN used in this model

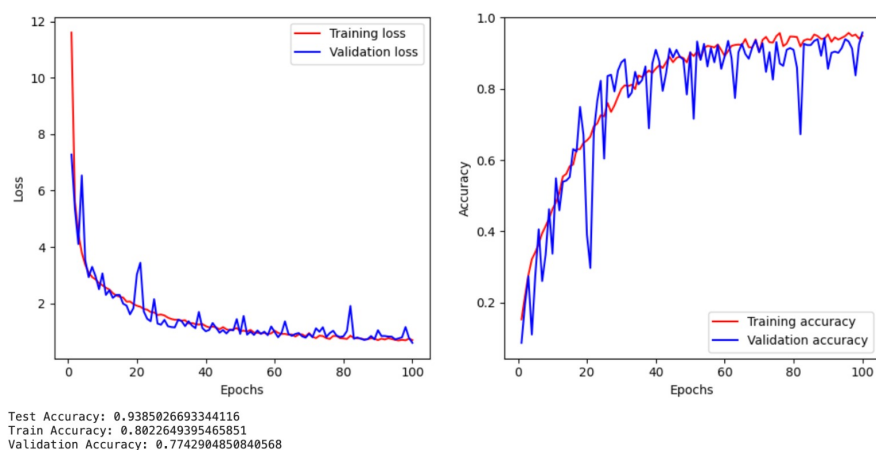


Figure 2: Plot illustrating training accuracy and validation loss