A picture containing drawing

Description automatically generated

**Python and Deep Learning Programming**

CSEE 5590 0001

Spring 2020

**Increment Project Report**

**Team Id - 1**

**Team Members:**

Sai Harshavardhan Maddula(class Id: 15 )

Venkata Sri Rohita Goparaju(class Id: 27)

Divya Reddy Bandari(class Id: 7)

Sai Srinivas Vidiyala(class Id: 17)

# Plant Seedlings Classification

# (Determine the species of a seedling from an image)

# [https://github.com/hvbilla/Python-Deep-Learning-/tree/master/project | You](https://github.com/hvbilla/Python-Deep-Learning-/tree/master/project%20|%20You) tube link

# Approach:

# In this project we are differentiating the weed from the crop seedling.

# The convolutional neural network (CNN) is a class of deep learning neural networks. CNNs represent a huge breakthrough in image recognition. They’re most commonly used to analyze visual imagery and are frequently working behind the scenes in image classification. They can be found at the core of everything from Facebook’s photo tagging to self-driving cars. They’re working hard behind the scenes in everything from healthcare to security.

# Image classification is the process of taking an input (like a picture) and outputting a class (like “cat”) or a probability that the input is a particular class (“there’s a 90% probability that this input is a cat”). You can look at a picture and know that you’re looking at a terrible shot of your own face by convolutional neural network.

# Convolutional Neural Network:

* starts with an input image
* applies many different filters to it to create a feature map
* applies a ReLU function to increase non-linearity
* applies a pooling layer to each feature map
* flattens the pooled images into one long vector.
* inputs the vector into a fully connected artificial neural network.
* processes the features through the network. The final fully connected layer provides the “voting” of the classes that we’re after.
* trains through forward propagation and backpropagation for many, many epochs. This repeats until we have a well-defined neural network with trained weights and feature detectors.

**Data Description:**

The dataset has about 4750 labeled images (1.73GB) showing plants of 12 different types. The goal is to classify correctly the species shown on the 794 images (91MB) of the test set. All images are quadratic but vary in size. We are provided with a training set and a test set of images of plant seedlings at various stages of grown. Each image has a filename that is its unique id. The dataset comprises 12 plant species. The goal of the competition is to create a classifier capable of determining a plant's species from a photo. The list of species is as follows:

* Black-grass
* Charlock
* Cleavers
* Common Chickweed
* Common wheat
* Fat Hen
* Loose Silky-bent
* Maize
* Scentless Mayweed
* Shepherds Purse
* Small-flowered Cranesbill
* Sugar beet

A picture containing food, photo, sitting, small

Description automatically generated

A picture containing bird

Description automatically generated

**Dataset Description:**

* train.csv - the training set, with plant species organized by folder
* test.csv - the test set, you need to predict the species of each image
* sample\_submission.csv - a sample submission file in the correct format

**Technical Stack:**

* Pycharm
* Python 3.8
* cv2
* glob
* numpy
* matplotlib
* math
* pandas

# System Architecture:

# 

**Project Implementation:**

**Step :1 Data Extraction:**

Data extraction is the first step of the project. As part of it, we need to import cv2, glob: which is used to retrieve the files matches with the specified pattern, numpy, matplotlib and pandas.

We are assigning 70 to scale variable and this variable is used to call all the pictures in our library to 70 pixels. All the images which end with .png extension are accredit to the path variable. We are retrieving all the images in the specified pattern in using glob function. Total count of the files will be assigned to the num.

All the images are resized using the loop function and assigned to the train dictionary.

Similarly, we are assigning all the labels of the images to the trainlabel.

Cleaning data:

As a part of cleaning the data. We need to separate the subject from background. Since all the photos of our plants in green colour, we are creating a mask to remove the background part of the image.

# Team Contribution:

Sai Harshavardhan Maddula : 25%

Venkata Sri Rohita Goparaju : 25%

Divya Reddy Bandari : 25%

Sai Srinivas Vidiyala : 25%

**Conclusion:**