KLE Society's

KLE Technological University

March 2, 2021

Course Project
On
V2 Plant Seedlings Dataset

Machine Learning(17ECSC306)

Team Number : 06
Submitted by

Name	Roll No	USN
Saloni Shah	363	01FE18BCS183
Samarth M	364	01FE18BCS184
Samarth R	365	01FE18BCS185
Sanjana Kambar	370	01FE18BCS190
Rajeshwari Hiremath	384	01FE18BCS167

Under The Guidance Of
Dr. Meena S M
Mr. Uday Nagraj Kulkarni

SCHOOL OF COMPUTER SCIENCE ENGINEERING
HUBLI-580031

Academic year 2020-2021

Abstract

Weed Control is a challenging problem these days, which is reducing the crop quality and productivity. Weeds are constantly competing for nutrients, water and sunlight. Thus, weed control must be achieved as early as possible, before the weeds begin to compete with the main crops and cause adverse effects.

With the motivation to develop a system to classify plant seedlings at their early growth stages, we propose a model for the classification of same. This approach extracts features from a given input image and classifies the image as class of a weed.

We have considered 12 classes of weeds namely, Common Chickweed, Maize, Black Grass, Sugar Beet, Scentless Mayweed, Small-flowered Cranesbill, Common Wheat, Cleavers, Charlock, Loose-Silky Bent, Fat Hen and Shepherd's Purse. The input image is classified as one of the mentioned classes.

Acknowledgement

Every project is successful largely due to the effort of a number of people who have always given their valuable advice or lent a helping hand. We sincerely appreciate the inspiration, support and guidance of all those people who have been instrumental in making this project a success.

We take this opportunity to thank Dr.Ashok Shettar, Vice-chancellor,KLE Technological University and to Dr. Prakash Tiwari, Principal, B V Bhoomaraddi College of Engineering and Technology, Hubli.

We take this opportunity to thank Dr.Meena S M, Head, School of Computer Science and Engineering for having provided us with an academic environment which nurtured our practical skills contributing to the success of our project.

We also take this opportunity to thank Prof. Uday Kulkarni for the guidance and constant supervision as well as for providing necessary information regarding the project also for supporting us throughout the project. We are thankful and fortunate enough to get constant encouragement, support and guidance from all teaching and non teaching staff of SoCSE which helped us in successfully completing our project.

We owe our gratitude to the School of Computer Science Engineering, KLE technological University for providing us with the resources necessary for the completion of this project.

Our gratitude will not complete without thanking the Almighty, our beloved parents, our seniors and our friends who have been a constant source of blessings and aspirations.

Contents

1 Introduction

1.1 Overview of the project

Weed control is a challenging problem that may face crop productivity. Weeds are perceived as an important problem because they conduce to reduce crop yields due to the expanding competition for nutrients, water, and sunlight besides they serve as hosts for diseases and pests. Thus, it is crucial to identify weeds in early growth in order to avoid their side effects on crops growth. The dataset used, provided by the Aarhus University Signal Processing group, in collaboration with University of Southern Denmark, contains a set of 5608 images of approximately 960 unique plants belonging to 12 species at several growth stages. The images are grouped into 12 classes. These classes represent common plant species in Danish agriculture. Each class contains rgb images that show plants at different growth stages. The images are in various sizes and are in png format.

1.2 Motivation

In order to improve agronomic production and crop quality, farmers should follow precision agriculture. Precision agriculture is a farm management approach that utilizes information technology and artificial intelligence to guarantee profit maximization, crop yield optimization, and environment preservation. One of the fundamental challenges that face precision agriculture is weed control. Weed control must be achieved earlier as possible after crop germination before weeds begin to compete with crops for nutrition and cause adverse effects. Thus, optimal weed treatment is recommended in the seedling stage. In accordance with it, this motivates us to develop a system to classify plant seedlings at their early growth stages.

1.3 Objectives:

- Identify weeds among the seedlings
- Reduce human efforts and saving time
- Increase plant yield

1.4 Literature survey

The Authors in [1] used deep learning to detect five tomato leaves diseases. They achieved a high accuracy in detecting the tomato disease. The authors in [2] provided a dataset that is aimed at ground-based weed or species spotting and also suggested a benchmark measure to researchers to enable easy comparison of classification results. The authors in [3] demonstrated the effectiveness of a convolutional neural network to learn unsupervised feature representations for 44 different plant species with high accuracy. In the course of exploring the right architecture for our model, we consider the work of [4] in classifying leaves using the VGGNet16 architectures. The authors in [5] implemented a 26-layer deep learning model consisting of 8 residual blocks in their classification of 10,000 images of 100 ornamental plant species achieving classification rates of up to 91.78 percent. The authors in [6] addressed the problem of CNN-based semantic segmentation of crop fields separating sugar beet plants, weeds, and background solely based on RGB data by proposing a deep encoder-decoder CNN for semantic segmentation that is fed with a 14-channel image storing vegetation indexes and other information that in the past has been used to solve crop-weed classification.

1.5 Problem definition

Successful cultivation of maize depends largely on the efficacy of weed control. Weed control during the first six to eight weeks after planting is crucial, because weeds compete vigorously with the crop for nutrients and water during this period. Annual yield losses occur as a result of weed infestations in cultivated crops. Crop yield losses that are attributable to weeds vary with type of weed, type of crop, and the environmental conditions involved. Generally, depending on the level of weed control practiced yield losses can vary from 10 to 100 percent. Rarely does one experience zero yield loss due to weeds. Yield losses occur as a result of weed interference with the crop's growth and development. This explains why effective weed control is imperative. In order to do effective control the first critical requirement is correct weed identification.

2 Approach

2.1 About the Project

In traditional image classification algorithms, handcrafted features are firstly extracted, then a feature selection process is achieved, and finally, a suitable classifier is chosen. However, CNN is proficient in learning various features from images, it covers global and local features, and it uses these features for efficient classification. The proposed system proceeds in four phases; preprocessing, constructing the network model architecture, training the network model and defining its parameters, and finally testing the designed network model.The convolutional neural network is randomly initialized, and then it is trained for performing the classification process and indicated a convolutional model. The weights of the CNN are updated utilizing the training set.For each iteration, the training and validation losses are computed.

2.2 Description of target users

The target users of the system will be the farmers and agriculturists, who will be looking forward to automated devices to identify weeds among the seedlings. Secondary target customers would be hobby planters, where they will depend on our system to distinguish any weeds from the seedlings.

2.3 Datasets

This dataset contains 5,539 images of crop and weed seedlings. The images are grouped into 12 classes as shown in the above pictures. These classes represent common plant species in Danish agriculture. Each class contains rgb images that show plants at different growth stages. The images are in various sizes and are in png format.

2.4 Advantages/Applications of proposed system

- Increases plant yield and plant life
- Reduce human efforts
- Saves time

3 Software Requirement Specification

3.1 Overview of SRS

The system will be deployed on a portable device. The portable device can be used to get the image of the seedling which will be classified as weed or not, and also will predict the type of weed/seedling.

3.2 Requirement Specifications

The system offers to classify/identify whether the seedling is a weed or not, and also predicts the name of the weed/seedling.

3.2.1 Functional Requirements

- Users shall be able to input the image of the seedling to be classified into the system.
- The system must correctly identify the seedling in the input image.

3.2.2 Non-Functional Requirements

- The system should perform identification of the seedling within 5 seconds.
- The system should not take more than 10 seconds to load.
- The system should be able to classify the image within 5 seconds after capturing the image.
- The accuracy of the system should be more than 80
- The system should be safe to use and not cause any harm to the user.
- The system should not damage the seedlings which are to be identified.
- The system should be easy to use by the end-user.
- The system should be portable.

3.3 Software and Hardware Requirement Specifications

The system needs a device with an operating system to run on. The device preferably needs to have a RAM and secondary storage. Hardware requirements necessary for the system would be a tripod to place the device on, to get a clear image.

4 Architecture

The input layer of the network expects a RGB image.The input image is passed through three convolutional blocks.Convolutional filters with a receptive field of 3 x 3 are used. Each convolutional block includes 2D convolution layer operation (the number of filters changes between blocks). All hidden layers are equipped with a ReLU (Rectified Linear Unit) as the activation function (nonlinearity operation) and include spatial pooling through use of a max-pooling layer followed by a dropout layer. The network is concluded with a classifier block consisting of two Fully Connected (FC) layers.

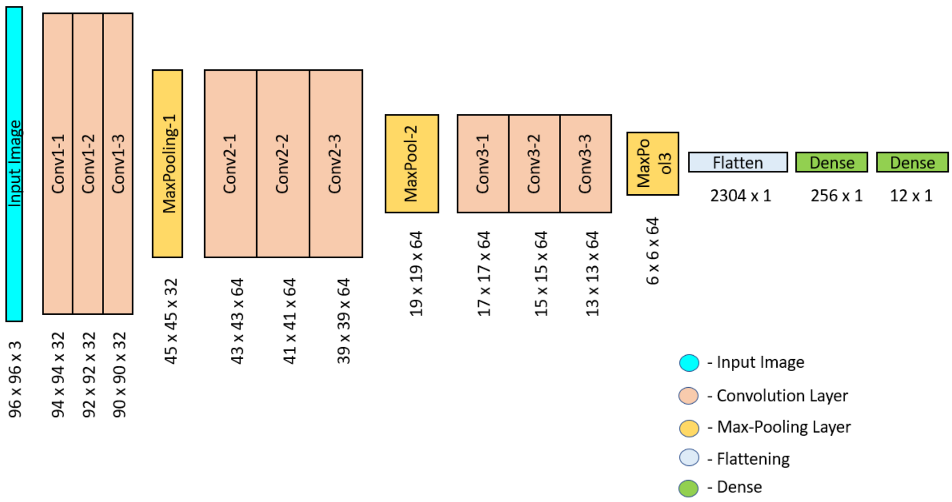

Figure 1: Pipeline

The input layer is fed with resized images (96 x 96 x 3). The hidden layer consists of 3 stages of learning layers. The utilized filters are all of kernel size 3x3 with a number of filters 32, 64 and 64 for respective convolutional blocks. The convolutional layers are associated with ReLU activation layers. Each convolutional layer is followed by a pooling layer and batch normalization layer.

4.1 Description of modules

4.1.1 Module 1: Data Augmentation

Input: Dataset
Output: Augmented dataset with all the classes having equal number of images. This is done so as to improve the training of the model.

4.1.2 Module 2: Data Image Generator

Input: Folder containing all the input images
Output: Dataset with all the images having class label as folder name.

4.1.3 Module 3: Simple Keras Model

Input: Images of seedlings of dimension 96x96 with 12 different classes
Output: Trained model which will identify the class of seedling

The configurations used are:

- OS: Windows 10
- GPU: NVIDIA GeForce GTX 1060
- CUDA TOOLKIT: v10.0
- cuDNN SDK: v7.5 (corresponding to CUDA TOOLKIT v10.0)
- Python: 3.x
- tensorflow-gpu: 1.13.1

5 Results and Discussions

The model being trained for 150 epochs using CrossEntropy loss and Adam optimizer along with a learning rate of 0.0001. The model has achieved:

- Train Accuracy : 0.9703
- Train Loss : 0.0867
- Validation Accuracy : 0.9515
- Validation Loss : 0.2188

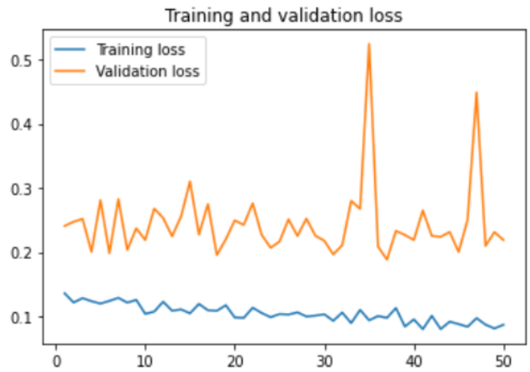

Figure 2: Training Loss - Validation Loss (For 50epochs)

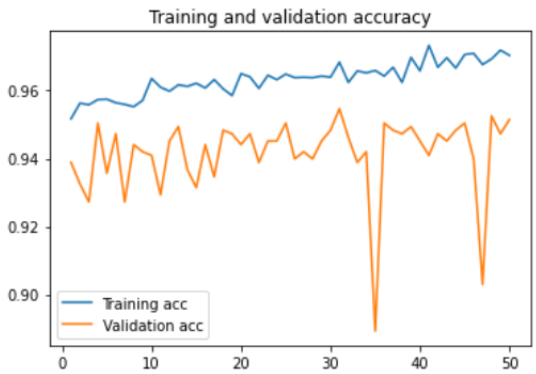

Figure 3: Training Accuracy - Validation Accuracy (For 50epochs)

6 Future Scope

- We proposed a solution for assisting farmers to optimize crops, which uses CNN model to classify the given 12 sets of seedlings into crops and weeds. The proposed approach achieves validation percent accuracy of 95.15.
- The future scope would be to develop an android/ios app which could be easily accessed by the farmers.

7 References

[1] Ashqar, B.AM, Abu-Naser, S.S(2019). ImageBased Tomato Leaves Diseases Detection Using Deep Learning. International Journal of Academic Engineering Research(IJAER) 2 (12), 10-16.

[2] T. Giselsson, R. Jørgensen, P. Jensen, M. Dyrmann, and H. Midtiby, A public image database for benchmark of plant seedling classification algorithms, CoRR, abs/1711.05458, 2017.

[3] Jeon, Wang-Su, and Sang-Yong Rhee, Plant leaf recognition using a convolution neural network, International Journal of Fuzzy Logic and Intelligent Systems 17, no. 1, pp 26-34. 2017.

[4] S. Lee, C. Chan, S. Mayo, and P. Remagnino, How deep learning extracts and learns leaf features for plant classification, Pattern Recognition, vol. 71, pp. 1-13, 2017.

[5] Y. Sun, Y. Liu, G. Wang, and H. Zhang, Deep learning for plant identification in natural environment, Computational Intelligence and Neuroscience, 2017.

[6] A. Milioto, P. Lottes, and C. Stachniss, Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs, cs.CV, 2018. URL