Arabidopsis and Tobacco Plant Classfication using Image Processing and Machine Learning.

Anant Krishna Mahale
School of Computer Science and Engineering
University of New South Wales
Sydney, Australia
z5277610@student.unsw.edu.au

I. Introduction

Botanists can easily identify and classify the plants, however for machines it's complex and computationally expensive. This part (Individual Component) of the project focuses on the, classification of the Arabidopsis and Tobacco plants using traditional computer vision techniques for feature extraction and pre-process the images. The data [1] has been captured in 2 separate laboratories from the top view images at once or many in the scene. It consists of Metadata.CSV which has information about genotype, treatment and time after germination and RGB images. There are 165 RGB images of Arabidopsis plant with height and width varying from 315 x 298 to 51 x 48 where as, 62 images of tobacco plants, with constant resolution 2448 x 2048.

The recent articles and technical papers focus on Deep learning approach, as it has improved the image recognition significantly compared to traditional image processing techniques. This project focuses on the traditional computer vision techniques as per the project specifications.

II. Literature Review

Anantrasirichai et al [5] proposed that Non-green backgrounds, such as soil, can be simply removed using an appropriate colour histogram threshold. An excess green index (ExG=2G-R-B) and an excess red index (ExR=1.4R-G-B), where R, G and B are red, green and blue of RGB colour space respectively, are computed. The difference between these indices is subsequently divided into two groups according to an Otsu threshold. If the are multiple paeks threshold is set one standard deviation below the mean.

Wu et al. [6] proposed an efficient algorithm for plant classification. They involved 32 kinds of plants. Several features such as aspect ratio (ratio between length and width of leaf), ratio of perimeter to diameter of leaf, and vein features were used to characterize the leaf with accuracy of `90.312%`

Chaki et al [7] proposed a new method of characterizing and recognizing plant leaves using a combination of shape features and texture. A Gabor filter was used to model the texture of the leaves and the shape was captured using a set of curvelet coefficients together with invariant moments. The efficiency of the system was tested using two neural classifiers which achieved a accuracy of `87.1%`

III. Methodologies

A. Foreground Extraction

The given image has different contrast and exposure resulting in skewed histogram. This can seen the Figure1. Considering the fact that all the channels are important, the histogram has to be equalized. Splitting the channels and equalizing each channel results in colour distortion. To address this problem, the RGB should be transformed into another space that contains a luminescence/intensity value (Luv, Lab, HSV, HSL) [3]. Equalizing the intensity plane and applying inverse transform results in better image. To retain the green part of the image and mask the background of the image [4], RGB colour-space is converted to HSV colour-space. Pixel values which fall outside HSV range of (45,54,66), (86, 255,255) are masked masked with black colour. Now the image has only plants (foreground). The results can be seen in the Figure 2

B. Feature Extraction using SIFT

The novel image feature extraction was described at 1999 by David G. Lowe who is researches from University of British Colombia. This method is named as Scale-Invariant feature Transform (SIFT) [8]. SIFT is a method that is invariant to scale, rotation, and illumination condition. Since the data-set has different size of leaves and orientation, sift was used to extract the features from the images.

Key locations in the image are defined as maxima and minima of the result of difference of Gaussians function applied in scale space to a series of smoothed and resampled images. Low-contrast candidate points and edge response points along an edge are discarded. Dominant orientations are assigned to localized keypoints. SIFT descriptors robust to local affine

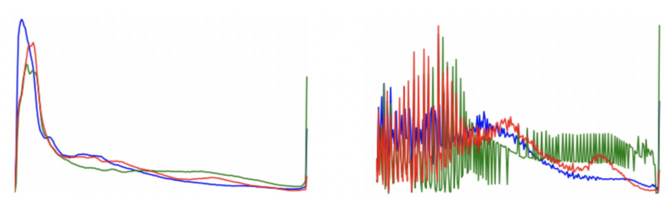

Fig. 1. Histogram of RGB Channels before (left) and after (right) equalization of V channel by converting into HSV Colour space and back

distortion are then obtained by considering pixels around a radius of the key location, blurring and resampling of local image orientation planes.

C. Feature Extraction using HOG

One of the popular method of feature extraction is Histogram of Oriented Gradients (HOG) [9]. In this method, an image is described by a set of local histograms. Then, the occurrences of gradient orientation is accumulated in a small spatial localized portions of the image referred as cell. The subsequent concatenation of 1-D histograms produces the features vector. Let the intensity value of the image to be analyzed is L. If the image is divided into N x N cells of size then the orientation $\theta x, y$ of the gradient in each pixel is calculated by using the Equation (1)

$$theta_{x,y} = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \tag{1}$$

The successive orientation belonging to the same cell j are quantized and accumulated into an M-bins histogram.

D. Bag of Visual Words

SIFT features have high dimensionality. For image classification, it is necessary to reduce the dimensionality of the feature space. Using clustering method on extracted features we can group them and build a vocabulary which intern is used to build the histogram of individual image. There are 4 stages involved while constructing Bag of Visual Words

- Extracting the features from the images using detector.
- Grouping descriptor to the set of clusters (vocabulary) with vector quantization algorithm using K-means.
- Construction of a bag of features, which calculates the number of features that are entered on each cluster
- Classification, training bag of features as feature vectors, and determine category of the image

E. SVM Classifier

Support Vector Machine (SVM) is popular classification method nowadays. It works well with high-dimensional data, and SVM can use kernel features to map original data to greater dimensionality. In contrast to another classification method, SVM not use all of data to be learned in the learning process, but just several chosen data is contributed to build model in learning process. This research use SVM because features used have big dimensionality depending on the number of vocabulary. The technical paper [10] was used as a reference.

IV. Experimental Setup

The experiments were conducted on the Google Collab in CPU runtime mode. Significant Python packages are Numpy, OpenCV, scikit-image, scikitlearn. The pre -processed images are kept in the Google Drive for feature extraction and training the model.

The dataset was utilized in 3 different ways.

- **Resizing the images without using any image processing techniques.**
  In this setting, a given image was resized to 200*200 and no other image processing technique was used.
- **Masking the background of resized 200*200 image**
  In this setting, the background was masked by extracting green colour in the range((45,54,66), (86, 255,255)) from HSV Colour Space.
- **Discarding the low-resolution Arabidopsis images**
  In this setting, only top 62 high resolution images of Arabidopsis plants with black background were selected.

Feature extraction using SIFT(Scale-Invariant Feature Transform)

SIFT implemented in non-free OpenCV module was utilized. Different parameters were tuned to refine the feature extraction. `nfeatures`: No value was specified as all the extracted features will be used for clustering. `edgeThreshold`: Since majority of the dataset had differentiable edges, value 10 was selected. `sigma`: Arabidopsis plants have blurry images, with the value 1.6, the algorithm was detecting more features on the plant. `nOctaveLayers`: This was set to default value as SIFT automatically calculates this value from the given image.

Feature extraction using HOG(Histogram of Oriented Gradients) HOG implemented in skimage was utilized. HOG performs slower compared to SIFT algorithm. Keeping in this mind, the hyper-parameters were tuned. `pixels_per_cell`: this value was set to 8 * 8 `cells_per_block`: was set to 2 * 2 keeping performance in mind. `feature_vector`: was set to true as further steps needs vector to train the machine learning model.

The obtained descriptors are grouped into set of clusters (vocabulary) with vector quantization algorithm using K-means with $k = 100$

The data was split into 80:20 ratio for training and testing purposes.

A SVM classifier was initialized with below parameters `gamma`: this defines the influence of each training example. A value of 0.1 was selected for gamma. `class_weight`: There is a class imbalance between both the classes. which is why class weight of 0.7 and 1.8 was assigned.

In order to understand the performance of the model, it is necessary to use different metrics. **Precision** measures how good is the model when the prediction is positive. **Recall** measures how good is the model predicting positive classes (In this case Tobacco)

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

For Precision and Recall **Micro Average** is considered.

Both of the above mentioned metrics do not consider both Positive and Negetive class. ROC curve summarizes the performance of the model at different threshold values by combining confusion matrices at all threshold values. X axis of ROC curve is the true positive rate (sensitivity) and y axis of the ROC curve is the false positive rate (1- specificity). AUC aggregates the performance of the model at all threshold values. The best possible value of AUC is 1 which indicates a perfect classifier.

Considering the fact that the length of the dataset (227) is very small it is necessary to employ re-sampling procedure to effectively evaluate the performance of the model. Cross Validation is one such re-sampling method that ensures every observation from the original dataset has a chance of appearing in training and test set. Hence K-Cross validation was performed with $k = 5$

V. Results and Discussion

As described in the experimentation section, the data was utilized in 3 different ways. For the convenience it is named `Data_Setting_*`. The Table I and II summarises the performance of SVM classifier on feature extracted using SIFT and HOG descriptors respectively. The Precision, Recall and AUC of both the experiments in all the `Data_Setting_*` performs identical to ideal / perfect model. The values remained same when K-Cross validation was performed with the value $k = 5$ in both the experiments.

TABLE I
Metric Values of SVC Classifier Trained using features extracted with SIFT Descriptor

	Precision	Recall	AUC
Data_Setting_1	0.98	0.94	1.0
Data_Setting_2	1.0	1.0	1.0
Data_Setting_3	1.0	1.0	1.0

TABLE II
Metric Values of SVC Classifier Trained using features extracted with HOG Descriptor

	Precision	Recall	AUC
Data_Setting_1	1.0	1.0	1.0
Data_Setting_2	1.0	1.0	1.0
Data_Setting_3	1.0	1.0	1.0

Both the leafs of the plants have different shape and are not identical in any way. As well, the images of Arabidopsis plants are obtained by cropping a bigger image where as each image of the Tobacco plant captured individually. Because of this difference in the process of capturing the images, tobacco images have steel tray in the background in contract Arabidopsis images which have mostly mud. Irrespective of how tuned SIFT descriptor is, there are few key-points that are detected on steel tray of the Tobacco plants. When a Dictionary of Visual Words are built using these key-points, tray on the Tobacco plant image becomes one of the significant features contributing to tobacco plant classification. Even though `Data_Setting_2`has black background, the steel tray in the tobacco images is still visible in many of the images. This can be seen in the Figure 3 They continue to act as potential key-points. This makes any classifier perform identical to ideal classifier.

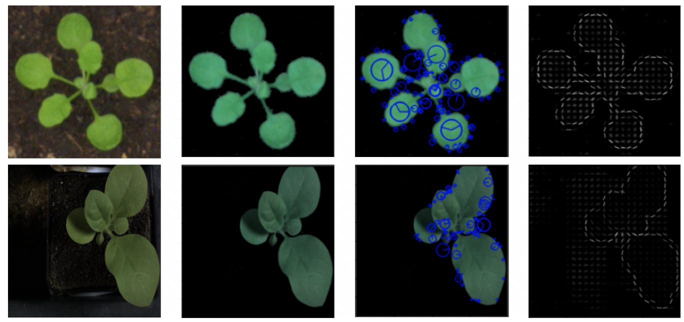

Fig. 2. Given Images (1), Masked Background (2), SIFT (3) and HOG (4) Descriptors

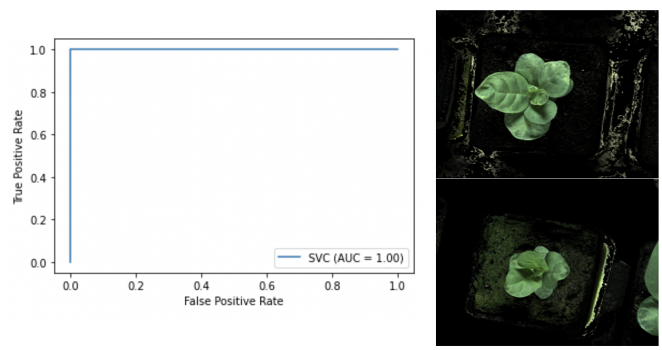

Fig. 3. AUC Curve (left) Images that weren't segmented correctly (right)

References

[1] M. Minervini, A. Fischbach, H.Scharr, and S.A. Tsaftaris. Finely-grained annotated datasets for image-based plant phenotyping. Pattern Recognition Letters, pages 1-10, 2015, doi:10.1016/j.patrec.2015.10.013

[2] website https://www.plant-phenotyping.org/datasets

[3] https://stackoverflow.com/questions/31998428/opencv-python-equalizehist-colored-image

[4] https://stackoverflow.com/questions/47483951/how-to-define-a-threshold-value-to-detect-only-green-colour-objects-in-an-image

[5] Anantrasirichai, Nantheera Hannuna, Sion and Canagarajah, Nishan. (2017). Automatic Leaf Extraction from Outdoor Images.

[6] S. G. Wu, F. S. Bao, E. Y Xu, Y-X. Wang, Y-F. Chang, Q-L. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", IEEE 7th Interantional Symposium on Signal Processing and Information Technology, Cairo, 2007.

[7] Chaki, J. et al. "Plant leaf recognition using texture and shape features with neural classifiers." Pattern Recognit. Lett. 58 (2015): 61-68.

[8] D. G. Lowe, "Distinctive Image Features From Scale-Invariant Key-points," pp. 1–29, 2004.

[9] Dalal, N. and B. Triggs. "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1 (June 2005), pp. 886–893.

[10] Christopher J.C. Burges, " A tutorial On Support Vector Machine For Pattern Recognition," Journal Data Mining and Knowledge Discovery, Vol.2 Issue 2, June 1998, pp 121-167.