

Automatic fruit classification using random forest algorithm

Hossam M. Zawbaa^{1,3,5}, Maryam Hazman^{2,5}, Mona Abbass^{2,5}, Aboul Ella Hassanien^{3,4,5}

¹Faculty of Mathematics and Computer Science, Babes-Bolyai University, Romania

²Central Lab. for Agricultural Expert System, Agricultural Research Center, Egypt

³Beni Suef University, Faculty of Computers and Information, Beni Suef, Egypt

⁴Cairo University, Faculty of Computers and Information, Cairo, Egypt

⁵Scientific Research Group in Egypt (SRGE)

<http://www.egyptscience.net>

Abstract—The aim of this paper is to develop an effective classification approach based on Random Forest (RF) algorithm. Three fruits; i.e., apples, Strawberry, and oranges were analysed and several features were extracted based on the fruits' shape, colour characteristics as well as Scale Invariant Feature Transform (SIFT). A preprocessing stages using image processing to prepare the fruit images dataset to reduce their color index is presented. The fruit image features is then extracted. Finally, the fruit classification process is adopted using random forests (RF), which is a recently developed machine learning algorithm. A regular digital camera was used to acquire the images, and all manipulations were performed in a MATLAB environment. Experiments were tested and evaluated using a series of experiments with 178 fruit images. It shows that Random Forest (RF) based algorithm provides better accuracy compared to the other well know machine learning techniques such as K-Nearest Neighborhood (K-NN) and Support Vector Machine (SVM) algorithms. Moreover, the system is capable of automatically recognize the fruit name with a high degree of accuracy.

Index Terms—Fruit classification, Image classification, Features extraction, Scale Invariant Feature Transform (SIFT), Random Forest (RF)

I. INTRODUCTION

Computer vision has been widely used in industries to aid in automatic and checking processes. Digital image processing plays an important role in the field of automation [1]. The important problem in computer vision and pattern recognition is the shape matching. It can be defined as the establishment of a similarity measure between shapes and its use for shape comparison. A result of recognition might also be a set of point identical between shapes. This problem has an important theoretical interesting motivation. Shape matching is intuitively accurate for humans, so there is a needed job which is not solved yet in its full generality. Shape matching applications contain image registration, object detection and recognition, and images content based retrieval [2].

Many agricultural applications used image processing to automate its duty. Detecting crop diseases is one of these applications. The crop images are analyzed to discovered the affected diseases [3]. Also, image processing are used

to monitor crop to decide its harvest time [4], [5], and to recognize fruits and vegetables type [6]. Recognize and classify fruits can aid in many real life applications. For examples, it can be used as an interactive learning tool to improve learning methods for children and as an alternative for manual barcodes in a supermarket checkout system [6], [7]. It could be helpful as an assist for the plant scientists to understand the genetic and molecular mechanisms of the fruits [7]. For eye weakness people, they can use it as a tool aiding them in shopping. Therefore, automation of fruit recognition and classification process is a big gain at both agriculture and industry fields, utilizing computer vision in food products has become very wide spread [8].

This paper goal is investigating the usage of Random Forest (RF) algorithm for developing an automatic fruit classifier. The proposed classification system includes preprocessing, features extraction, and classification stages. The fruit images feature is extracted based on the fruits' shape, colour characteristics and Scale Invariant Feature Transform (SIFT). The classification is done using Random Forest (RF) algorithm. The proposed approach is evaluated using a series of experiments with 178 fruit images and compared the Random Forest (RF) results with K-Nearest Neighborhood (K-NN) and Support Vector Machine (SVM) algorithms.

The structure of this paper is as follows. Section II describes the stages of the classification proposed system. Feature extraction and classification stages will be discussed in sections III and IV. Section V introduces the results of the experimental. Finally, the conclusion and future work is presented in section VI.

II. THE PROPOSED AUTOMATIC FRUITS CLASSIFICATION SYSTEM

This research aims to develop an automatic fruits classification system that recognizes a fruits image name from a collection of images. Achieving this aimed is done by proposing classification system which includes the following stages:

- Preprocessing Stage: in this stage, the images are resized to 90 x 90 pixels to reduce their color index. Also, the proposed model assumes that the acquired image is in RGB format which is a real color format for an image [9].
- Feature Extraction Stage: two feature extraction methods are used in this stage. The first one extracts the shape and color features. While, the second feature extract method uses the Scale Invariant Feature Transform (SIFT).
- Classification Stage: our proposed model used the Random Forests (RF) algorithm to classify the fruit image in order to recognize its name.

More detail about feature extraction and classification stages is given in the following sections.

III. FEATURE EXTRACTION STAGE

The aim of this stage is to extract the attributes or characteristics that describe an image. The classification accuracy are mainly depends upon feature extraction stage. So the presented work investigates using two methods for extracting images features which are shape and color features and Scale Invariant Feature Transform (SIFT).

Since, color consider as an significant feature for image representation due to the color is invariance with respect to image translation, scaling, and rotation [10]. Therefore, the first feature extraction method uses color and shape characteristics to generate the feature vector for each fruit image in the dataset. The used color moments to describe the images are color variance, color mean, color kurtosis, and color skewness [10], [11]. The shape features is described using Eccentricity, Centroid, and Euler Number features [12]. Eccentricity computes the aspect ratio of the distance of major axis to the distance of minor axis. It is calculated by minimum bounding rectangle method or principal axes method. The shape centroid defines the centroid position of image which is fixed in relative to the shape. The image Euler number defines relation between the connecting parts number and the holes number on image shape. Euler number is calculated by subtract the shape holes number from the contiguous parts number. [12].

The second method generates the feature vector uses the Scale Invariant Feature Transform (SIFT) algorithm [13]–[15]. It is an algorithm for image features extraction which is invariant to image rotation, scaling, and translation and partially invariant to affine projection and illumination changes. SIFT contains four main steps namely: scale-space extreme detection, keypoint localization, orientation assignment and keypoint descriptor [16]. The scale-space extreme detection step identifies the points of potential interest using difference-of-Gaussian function (DOG). In the keypoint localization, a model is fit to define location and scale for each candidate location. The selected Keypoints are determined based on their stability measures. In the orientation assignment step, orientations are allocated for each keypoint location according to the local image gradient directions. Then the operations are

- 1: Build the image Gaussian pyramid $L(m, n, \sigma)$ using the following equations 1, 2, and 3.

$$G(m, n, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(m^2+n^2)}{2\sigma^2}\right), \quad (1)$$

$$L(m, n, \sigma) = G(m, n, \sigma) * I(m, n), \quad (2)$$

$$D(m, n, \sigma) = L(m, n, k\sigma) - L(m, n, \sigma), \quad (3)$$

Where σ is the scale parameter, $G(m, n, \sigma)$ is Gaussian filter, $I(m, n)$ is smoothing filter, $L(m, n, \sigma)$ is Gaussian pyramid, and $D(m, n, \sigma)$ is difference of Gaussian (DoG).

- 2: Calculate the Hessian matrix.
- 3: After that, calculate the determinant of the Hessian matrix as shown in the equation 4 and eliminate the weak keypoints.

$$Det(H) = I_{mm}(m, \sigma)I_{nn}(m, \sigma) - (I_{mn}(m, \sigma))^2 \quad (4)$$

- 4: Calculate the gradient magnitude and orientation as in equations 5 and 6.

$$Mag(m, n) = ((I(m+1, n) - I(m-1, n))^2 + (I(m, n+1) - I(m, n-1))^2)^{1/2} \quad (5)$$

$$\theta(m, n) = \tan^{-1}\left(\frac{I(m, n+1) - I(m, n-1)}{I(m+1, n) - I(m-1, n)}\right) \quad (6)$$

- 5: Apply the sparse coding feature based on SIFT descriptors as in equations 7 and 8.

$$\min \sum_{i=1}^S (\|m_i - \sum_{j=1}^Z a_i^{(j)} \phi^{(j)}\|^2 + L) \quad (7)$$

$$L = \lambda \sum_{j=1}^Z |a_i^{(j)}| \quad (8)$$

Where m_i is the SIFT descriptors feature, a^j is mostly zero (sparse), ϕ is the basis of sparse coding, λ is the weights vector.

Algorithm 1: SIFT Feature Extraction Algorithm

done on image data which has been converted relative to the allocated scale, orientation, and location for each feature, so providing invariance to these transformations. In the keypoint descriptors step, the local image gradients are computed for the selected scale in a certain neighborhood around the identified keypoint. These are converted to a representation to allow for important levels of local shape distortion and modify in illumination [17]. It works according to Algorithm 1.

IV. CLASSIFICATION STAGE

In the classification stage, the proposed model applies the Random Forests (RF) classifier to recognize different kinds of fruits, and comparing the Random Forest (RF) results with K-Nearest Neighborhood (K-NN), Support Vector Machine

Download English Version:

<https://daneshyari.com/en/article/6664876>

Download Persian Version:

<https://daneshyari.com/article/6664876>

[Daneshyari.com](https://daneshyari.com)